2009

The effect of C8-arylguanine adducts on B/Z-DNA equilibrium: Implications in aryl hydrazine carcinogenesis

Vorasit Vongsutilers
West Virginia University

The Effect of C8-Arylguanine Adducts on B/Z-DNA Equilibrium: Implications in Aryl Hydrazine Carcinogenesis

Vorasit Vongsutilers

Dissertation submitted to the
School of Pharmacy
at West Virginia University
in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy
in
Pharmaceutical and Pharmacological Sciences

Peter M. Gannett, Ph.D., Chair
Patrick S. Callery, Ph.D.
Robert R. Griffith, Ph.D.
Yongyut Rojanasakul, Ph.D.
Björn Söderberg, Ph.D.

Department of Basic Pharmaceutical Sciences

Morgantown, West Virginia
2009

Keywords: Z-DNA, C8-arylguanine adduct, CD, NMR, aryl hydrazines, carcinogenesis

ABSTRACT

The Effect of C8-Arylguanine Adducts on B/Z-DNA Equilibrium: Implications in Aryl Hydrazine Carcinogenesis

Vorasit Vongsutilers

Aryl hydrazines and related compounds have been shown to be carcinogenic but the mechanism is still unclear. C8-Arylguanine adducts, formed from oxidative metabolites of aryl hydrazines, are suspected to be the cause of carcinogenesis. Z-DNA formation facilitated by the aryl adduct are among the potential mechanisms and has been investigated in this study. Z-DNA may be involved in carcinogenesis as several studies have indicated that Z-DNA may play an important role in gene expression and to induce mutagenic genetic deletions. C8-Arylguanine adducts may cause carcinogenesis by promoting the formation of Z-DNA which in turn disturbing gene regulation and leading to mutagenesis. To investigate this hypothesis, we have set out to 1) study the effect of aryl adducts formed from carcinogenic aryl hydrazines on B/Z-DNA equilibrium and 2) seek the relevancy between Z-DNA formation and aryl hydrazine carcinogenesis.

Here, alternating CG decamers containing C8-arylguanine modifications ($d(CGCGCG^*CGCG)_2$, G* = C8-tolyl, C8-carboxyphenyl, C8-methoxymethylphenyl, or C8-hydroxymethylphenyl guanine), were prepared through Suzuki coupling and phosphoramidite chemistry. The effect of the aryl adducts on the B/Z-DNA equilibrium were determined by CD and NMR analysis. The experimental results supported by computational study have suggested that all of the aryl modifications examined facilitate B-Z transition by destabilizing the B conformation and/or stabilizing the Z conformation relative to the corresponding unmodified oligonucleotide. Among the aryl adducts, C8-carboxyphenyl adduct has been shown to be best at facilitating B-Z transition followed by C8-phenyl, C8-methoxymethylphenyl, C8-hydroxymethylphenyl, and C8-tolyl respectively. The effect of aryl adducts on B-Z transition is generally correlated with the reported V79 mutagenicity of the aryl hydrazines precursors. This preliminary study has suggested that Z-DNA may be involved or plays a role in aryl hydrazine carcinogenesis.

ACKNOWLEDGEMENT

First of all, I would like to express my thanks to Dr. Peter M. Gannett, my advisor, for his guidance during my journey as a graduate student. His remarkable knowledge in various fields and more importantly, his understanding and encouragement, have greatly helped me get through hardship in my study. In addition, I would like to express my gratitude to my academic committee members Dr. Patrick S. Callery, Dr. Robert R. Griffith, Dr. Yongyut Rojanasakul, and Dr. Björn Söderberg for their many valuable suggestions and time to serve on my committee.

I would like to thank my colleagues in BPS department, especially in Dr. Gannett's lab, who share similar difficulties and remind me that everybody needs to face obstacles in work and life, encouraging me to keep going. Also, I would like to thank undergraduate and intern students who have worked with me and helping in several perspectives of this study. Many thanks go to staff in School of Pharmacy for their kind help during my time, here, in the graduate program.

Finally, to my family who stand by me no matter what, I could not possibly be successful without their help. I sincerely appreciate their lifetime, unconditional love and support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SCHEMES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

1.1 Carcinogenicity of Aryl Hydrazines and Related Compound

Aryl hydrazines and related compounds can be found in a variety of sources including natural products, pharmaceutical agents, synthetic chemicals used in industry, and agricultural substances[1] (Figure 1.1). *Agaricus bisporus*, the most produced mushroom in the United States, contains Agaritine which was found to be carcinogenic in mice after bioactivation[2,3]. Several medicines, such as isoniazid and hydralazine, contain aryl hydrazines or related functional groups. Though they display a wide range of carcinogenic activities, most of the aryl hydrazines that have been studied are carcinogenic[4]. However, though the carcinogenicity of aryl hydrazines has being investigated for more than thirty years, the mechanism(s) with respect to carcinogenesis remains unclear and is still being investigated.

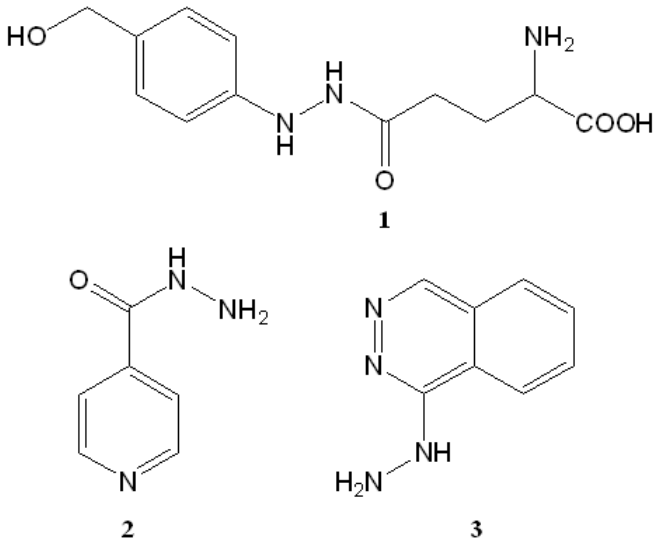


Figure 1.1 Structure of aryl hydrazines and related compounds found in everyday life (1) agaritine from the button mushroom *Agaricus bisporus*, (2) antitubercular agent isoniazid, and (3) antihypertensive hydralazine.

Oxidative metabolism of aryl hydrazines is believed to produce reactive intermediates that are related to the carcinogenicity of aryl hydrazines. Metabolism of aryl hydrazines by cytochrome P450 has been shown to lead to the production of the reactive intermediate arenediazonium ions, as first demonstrated by trapping with 2-naphthol[5,6]. *In vivo*, arenediazonium ions may react directly or, after reduction, to reactive aryl radicals (Scheme 1.1).

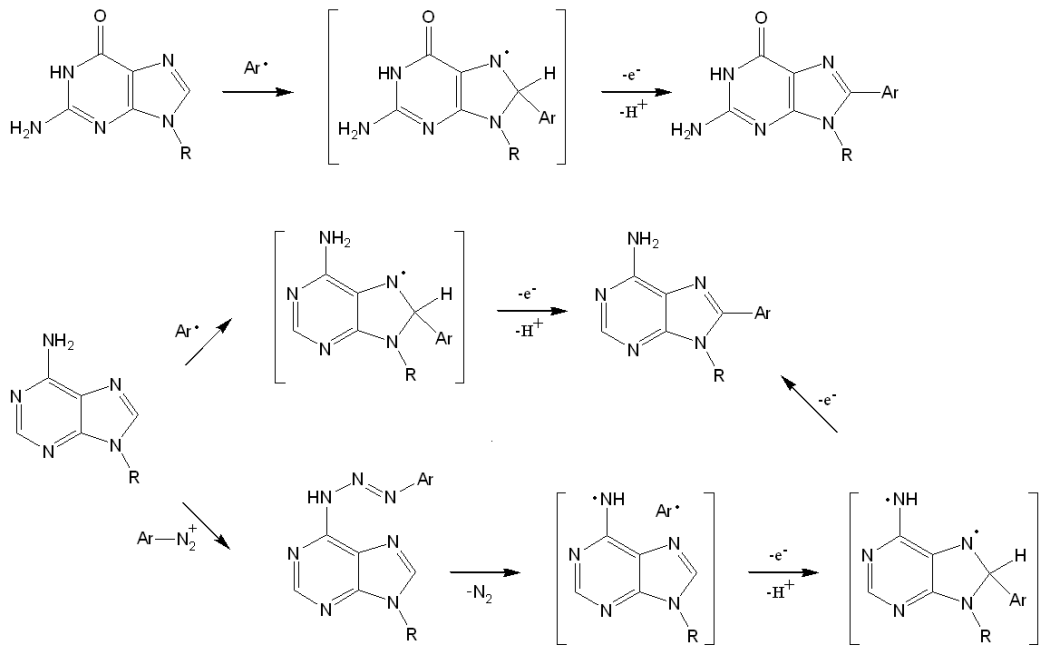
Scheme 1.1 Metabolism of aryl hydrazine leads to formation of arenediazonium ion and aryl radical.

Both arenediazonium ions and aryl radicals are reactive electrophiles and they can react with DNA either directly or indirectly. Direct reaction of the aryl radical leads to the formation of DNA adducts. Reaction with water or oxygen results in the formation of reactive oxygen species (ROS) which may, in turn, react with DNA. ROS can be produced by both chemical processes[7] and during enzymatic metabolism of aryl hydrazines, as shown for example, when methylphenyl hydrazines are incubated with microsomes from C50 cells[7,8].

Several ROS intermediates have been implicated in genetic alterations and carcinogenesis through oxidative DNA damage[9-11]. 8-Hydroxyguanine, also known as 8-oxoguanine, is one of the products resulting from ROS production and subsequesnt reaction with DNA, ultimately leading to base mis-pairs as Cheng K.C. et al. have

shown[12]. Generation of ROS, or the related reactive nitrogen specied (RNS) through metabolism may be a part of mechanism behind aryl hydrazine carcinogenesis. However, previous studies have not provided sufficient results to indicate a direct relationship between aryl hydrazines carcinogenesis and ROS/RNS production. Runge-Morris et al. suggested that hydrazine-mediated DNA damage is more likely to occur through organic free radical (carbon-centered) than from ROS[13]. In turn, the importance of direct DNA adducts formation has been increased, significantly, due to several studies that indicate the possibility of DNA adducts can alter either local or global DNA conformation.

The C8 position of the purine bases has been shown to be a common target for radical species such as aryl radicals formed from aryl hydrazines. As shown in Scheme 1.2, adenine can react with both the arenediazonium ion and aryl radicals to form a C8-aryl adenine adduct while C8-aryl guanine adduct formation generally occurs by reaction with the aryl radical directly[14]. Aryl radicals can react at the C8 position of purine bases to form carbon centered radical intermediates. This intermediate, then, must undergo an oxidation and loss of a proton to produce the neutral adduct. In the case of adenine, reaction at C6-N can also occur which yields a triazene. The triazene is unstable and can decompose, with the loss of nitrogen and the resulting aryl radical can react at C8. Tautomerization then leads to the final C8-aryl purine adduct.



Scheme 1.2 Mechanisms of C8-arylpurine adduct formation: The upper scheme shows C8-arylguanine adduct formation while the lower scheme shows C8-aryladenine adduct formation.

DNA adducts have been extensively investigated for their potential to cause cancer development. Miscoding[15] and DNA strand breakage[16] have been studied in relation to C8-aryl purine adducts, however none of these processes has unequivocally been related to aryl hydrazine carcinogenesis. Study of DNA adducts has often focused on the relationship of the observed conformational changes and the resulting genetic mutations. Of the many known C8 guanine adducts, the 2-aminofluorene[17], bromine[18,19], and methyl[20,21] adducts have all been shown to shift the B/Z equilibrium toward Z-DNA in alternating purine-pyrimidine sequences, along with the phenyl[22] adduct formed from aryl hydrazines. The effects of C8-aryl adduct on Z-DNA and B-DNA were examined throughout this study in order to evaluate the significance of the B-Z conversion in aryl hydrazine carcinogenesis.

1.2 Z-DNA: Structure and Biological Role

1.2.1 Structure and Chemistry of Z-DNA

Z-DNA was first identified by X-ray crystallography in 1979 by Wang et. al.[23]. The X-ray crystal structure indicated that a self complementary CG hexamer was in a non-canonical, left-handed double helix as shown in Figure 1.2. Although both B and Z-DNA are anti-parallel, double stranded helices that are composed of Watson-Crick hydrogen bonding base-pairs, their structures are very different. Unlike B-DNA in which all of the bases are in *anti* conformation about the glycosidic bond, in Z-DNA the pyrimidine bases are *anti* and the purines adopt the *syn* conformation (Figure 1.3). This structural feature leads to a zigzag phosphate backbone of the left-handed DNA which, in turn, gave rise to the name Z-DNA. Studies have shown that purine bases are more prone to adopt the *syn* conformation than pyrimidine bases mainly due to the energy penalty from van der Waals crowding[24]. Therefore, alternating purine-pyrimidine sequences are the most prone to form Z-DNA. In addition, Z-DNA does not have as visually distinct major groove and minor groove as seen in B-DNA. It does have a very narrow and deep minor groove while; what is technically the major groove, is nearly flat or even convex. The positions of the base-pairs, within the helix, are also significantly displaced from the helical axis in Z-DNA relative to the B form. This reduces stacking interactions among bases[25] and lowers the stability of the left-handed DNA form. Under standard physiological conditions B-DNA is typically lower in energy, though certain factors can stabilize Z-DNA relative to B-DNA and thus facilitate B to Z conversion.

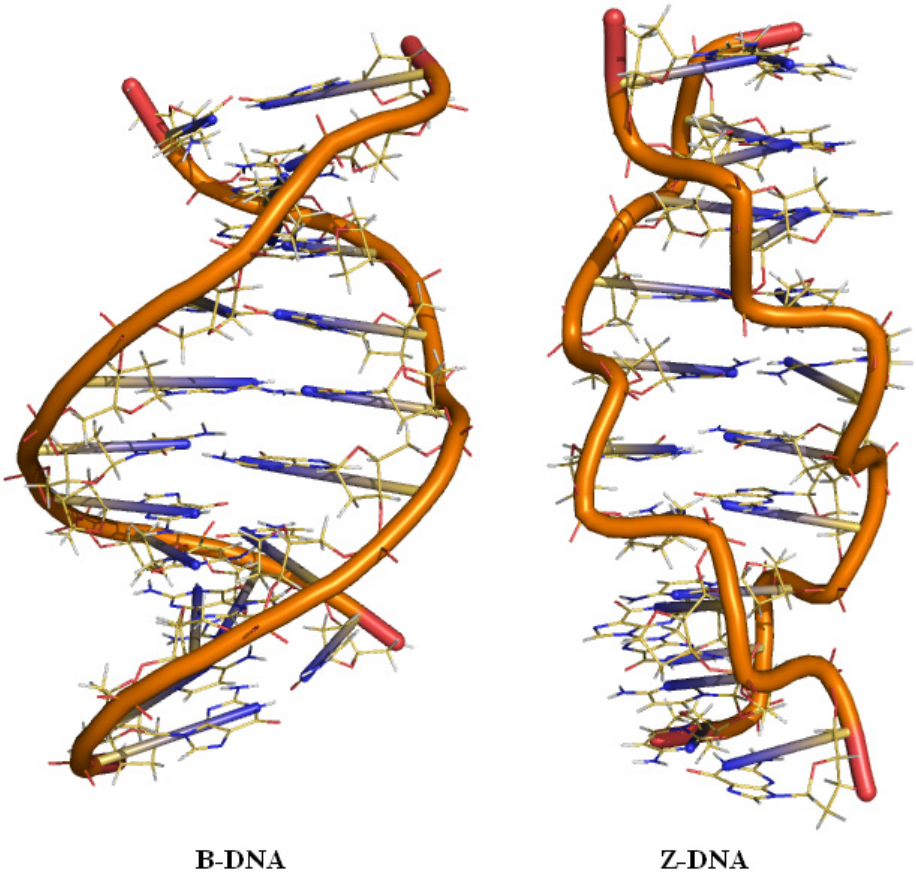


Figure 1.2 Structure of B-DNA (left) and Z-DNA (right) of the duplex CG decamers

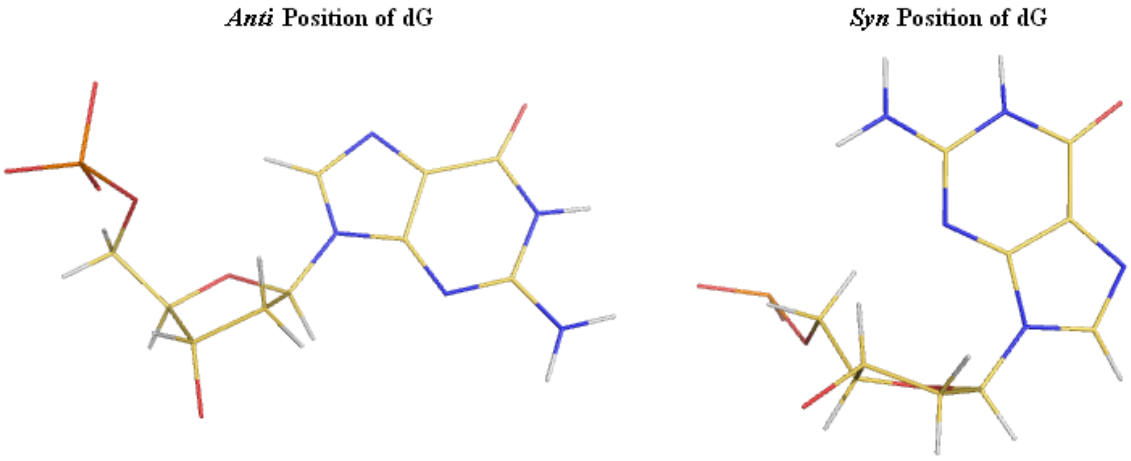


Figure 1.3 Structure of guanosine in B-DNA and Z-DNA. In B-DNA, the glycosidic bond in is in *anti* conformation and the deoxyribose adopts the C2′ endo orientation. In Z-DNA, the glycosidic bond is in *syn* and the deoxyribose contains is C3′ endo.

Several studies have investigated the factors affecting the stability of Z-DNA. Alternating purine-pyrimidine sequences are more common to form Z-DNA due to the previously described preferences for purines to adopt *syn* glycosidic bonds. Certain chemical modifications like N7 guanine methylation[26] or C5 cytosine methylation,[27] help stabilize Z-DNA and/or destabilize B-DNA and drive the B to Z conversion. A very interesting target is the C8 position of guanine since several modifications including methylation, bromination, and phenylation, have been shown to facilitate B-Z transition via C8 guanine adduct formation. The addition of a group to the C8 of guanine or adenine causes the *syn* conformation to be favored, at least at the nucleoside level, due to less steric hindrance compared to *anti* conformation. In *anti* position, the C8-substituent sterically interacts with the C2′-proton.

Organic and inorganic cations can stabilize Z-DNA and molecules with those bearing multiple positive charges are even more efficient at stabilization of the Z-DNA conformation than monovalent cations as has been shown in poly GC sequences[28]. The reason for the stabilizating effect of cations is, in Z-DNA, the phosphate groups are closer together than in B-DNA and under normal physiological conditions give rise to electrostatic repulsion among the negatively charged phosphate backbone. Positive charged species such as Na^+, K^+, Mg^+, spermine, and spermidine diminish the electrostatic effect by screening and thereby lower the energy level of Z-DNA.

In 1982, another factor that stabilizes Z-DNA was discovered. Negative supercoiling, which occurs during transcription, has been shown to facilitate B to Z conversion. This factor was originally detected with gel electrophoresis[29] and a Z-DNA antibody binding assay[30]. The level of negative supercoiling required to stabilized Z-

DNA depends on the length of alternating purine-pyrimidine track. Generally longer tracks of purine-pyrimidine repeats require a lower level of negative supercoiling to stabilize the Z conformation[25]. The discovery of the stabilization effect that negative supercoiling has on the Z-DNA conformation suggested possible biological roles for Z-DNA since supercoiling tends to be interrelated with biological function (e.g., transcription). Further details regarding supercoiling and biological significant of Z-DNA are discussed in the following section.

1.2.2 Biological Relevance of Z-DNA

Although many chemical aspects of left-handed DNA have been investigated and elucidated, the biological role has not. A major impediment is the lack of data regarding the formation of Z-DNA under physiological conditions, even after 30 years of research following the discovery of Z-DNA. However, in the past decade several studies have shown that Z-DNA may play significant roles in biological systems and investigations into its biological functions are being revived. The evidence that suggested that negative supercoiling stabilizes Z-DNA has redirected biologists to have renewed interest in Z-DNA function[31].

In 1987, study of DNA supercoiling during transcription by Liu L.F. and Wang J.C. [32] showed the possible biological relevance of Z-DNA in gene transcription. The study showed that, in the process of transcription, RNA polymerase actually plows through DNA helices instead of rotating around DNA, as believed in the past. Because the DNA has a fixed end, the movement of RNA polymerase complex unwinds the DNA and generates a negative supercoil behind the moving polymerase. The left-handed Z-

DNA form, then, may form to reduce the negative torsional strain caused by the unwound DNA. The information from this study supports the ideas that Z-DNA may have a biological function(s).

A computational study was conducted to map the potential Z-DNA forming sequences in 137 human genes[33]. The result has shown that potential Z-DNA forming sequences are located non-randomly near transcription initiation sites. Further evidence from a more recent computational study in 2004[34] also suggests that Z-DNA coupled transcription may be possible since the results show the distribution of Z-DNA forming regions across chromosome 22 toward the transcription start sites. To gain more direct evidence of Z-DNA associated with transcription, the level of Z-DNA during transcription in permeabilized mammalian cell nuclei was detected by a Z-DNA antibody binding assay[35]. This study has discovered that transcription is directly associated with the binding of anti Z-DNA antibodies; inhibition of RNA transcription results in decreasing Z-DNA antibody binding.

Specific gene expression, for example, of *c-myc*, has also been shown to correlate with the formation of Z-DNA in permeabilized nuclei[36]. In 2001, Liu R. et al. proposed one possible mechanism to explain how Z-DNA initiates transcription[37]. To activate the colony stimulating factor-1 (CSF-1) promoter, the activity of the NFI-BAF complex required the transition of B-DNA to Z-DNA. Binding of NFI to CSF-1 facilitates the activation of the promoter by BAF which in turn initiates the unwinding of chromatin structure. Negative supercoiling would then be increased and stabilized by the transition of B to Z-DNA. This discovery has shown that Z-DNA could not form nucleosomes[38]. Therefore, stabilization of Z-DNA, TG repeats in this case, by negative supercoiling

further opens the chromatin structure and allows other transcription factors to engage and to initiate transcription. Z-DNA formation is only transient and will relax back to B-DNA after transcription ends. The evidence of Z-DNA coupled CSF-1 transcription has demonstrated a significant biological role of Z-DNA.

In addition to Z-DNA coupled transcription, other studies have found the occurrence of Z-DNA binding proteins. A screening technique to identify selective Z-DNA binding proteins has been described by Herbert et al.[39] It led to the isolation of several specific Z-DNA binding proteins. One of the Z-DNA binding proteins that is worthy of mention is double stranded RNA adenosine deaminase (ADAR1)[40]. ADAR1 binds to the double stranded part of RNA that forms in pre-mRNA[31] and activates deamination of adenosine to give inosine, interpreted as guanine by ribosome. The activity of ADAR1 may regulate the expression of multiple proteins from single encoded genes through the alteration of amino acid sequence of the encoded protein. Although the *in vivo* experiment on Z-DNA binding domain of ADAR1 has been studied[41] and the co-crystallization of Z-DNA and ADAR1 has been conducted[42], a regulatory role of Z-DNA on ADAR1 function is still not completely elucidated.

1.2.3 Z-DNA Induces Genetic Instability

From previous studies, the biological relevance of Z-DNA has been demonstrated mainly in a transcription process. Since Z-DNA seems to have a biological role in gene expression, it has been speculated that Z-DNA may play a role in carcinogenesis. Considering the suggestion that Z-DNA may regulate transcription, over-stabilization of the Z-DNA conformation may have negative consequences. As explained earlier that Z-

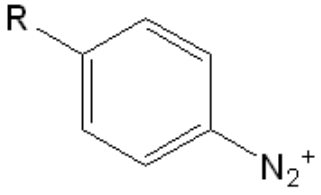
DNA formation only occurs transiently, initiating transcription and over stabilized Z-DNA may prolong Z-DNA formation and therefore, potentially over-expression of a gene. In addition, increasing the stability of Z-DNA may affect the regulatory role of Z-DNA on ADAR1 activity and can result in translation errors from improper guanine substitution. Furthermore, Z-DNA prone sequences have been shown to be highly recombinogenic and often result in genetic deletions[43]. In 1989, Spitzner, J. R. et al. [44] demonstrated that alternating purine-pyrimidine repeats were sensitive to strand breaking by DNA topoisomerases. The study suggested that the Z-DNA forming sequence may be a hot spot that is vulnerable to DNA damage. Recently, Wang G. et al[45] have confirmed that Z-DNA may cause genetic instability. They demonstrated Z-DNA is a genomic hot spot that can induce genetic instability in both bacterial and mammalian cells. Thus, a Z-DNA prone sequence induced double strand breaks close to the Z-DNA and result in large scale deletions. The result from this study has again suggested the mutagenic potential of Z-DNA. Considering that several C8 guanine adducts including C8-phenyl adduct formed from phenyl hydrazine[22] can stabilize the left-handed DNA, Z-DNA stabilization may be related to aryl hydrazine carcinogenesis.

1.3 Effect of C8-Arylguanine Adduct on Aryl Hydrazine Carcinogenesis: The Possibility of Z-DNA Mediated Carcinogenesis

Previous studies have shown a correlation between C8-arylguanine adduct formation and carcinogenicity of the parent aryl hydrazines[22,46]. Thus, through bioactivation of aryl hydrazines, arenediazonium ions and/or aryl radicals are produced and lead to the formation of DNA aryl adducts. This process may be the cause of aryl

hydrazine carcinogenesis. Several attempts have been made to determine how the DNA aryl adduct facilitates cancer development. Kohda et al.[15] has suggested that the C8-phenyl guanine adduct may be mis-read by DNA polymerase and can cause miscoding through G→T and G→C transversions. However, the C8-arylguanine adducts are not efficiently read and when read, tend to be read correctly. Consequently, aryl hydrazine carcinogenesis is likely caused by other mechanisms[22].

Four different *p*-substituted arenediazonium ions (Figure 1.4) including MBD, HMBD, MMBD, and CBD have been shown to be capable of inducing DNA damage through DNA-DNA cross linking (except for MBD) and single strand breaks in Chinese hamster lung fibroblast V79 cells[46]. It has been suggested that either the reactive azo group of arenediazonium ions or, after reduction to an aryl radical (metabolically or chemically), reaction with DNA occurs to form the C8-arylguanine adduct. Subsequently, *p*-substituents (-CH_2OH, -CH_2OCH_3, -COOH, CH_3) may then react with another base to form DNA cross linking[47]. The methyl functional group in MBD does not bear any leaving groups and, therefore, no reactions with other bases are possible and no DNA cross links form. The mechanism of DNA single strand breaks caused by reaction of the arenediazonium ions or aryl radicals is still not clear and will require additional research to elucidate the mechanism.



a) R = -CH_3
b) R = -CH_2OH
c) R = -CH_2OCH_3
d) R = -COOH

Figure 1.4 Structures of carcinogenic arenediazonium ions a) MBD, b) HMBD, c) MMBD, and d) CBD.

Studies of carcinogenic polycyclic aromatic hydrocarbons have demonstrated that the C8-aryl modified guanosine formation through the oxidation of benzo[α]pyrene can result in the loss of the pendant sugar residue[48]. Identification of depurination in both *in vitro* and *in vivo* studies[49,50] have been conducted and the results suggest that the C8-benzo[α]pyrene adduct formation facilitates the depurination of the modified guanine. Furthermore, the formation of apurinic sites has been correlated with the mutagenic level of benzo[α]pyrenes[49]. Depurination is a common occurance in cells and is usually taken care by DNA repair enzymes[51,52]. However, an excessive exposure to certain types of chemicals, for example carcinogens, may overwhelm the DNA repair enzymes. In this case, the abasic sites may cause several types of DNA lesions and may lead to mutagenesis[51]. DNA strand breaks can occur through the abasic sites since the aldehydic lesions are vulnerable to oxidative strand scission[53]. Misincorporation by DNA polymerase through depurination can occur as explained by Boiteux and Laval's observation[54] that *E. coli* DNA polymerase preferentially incorporates dAMP over dGMP opposite the abasic site. In addition, Dutta et al.[55] have shown that the abasic sites can generate interstrand cross links through the reaction of the aldehyde of the abasic site and N2-amino group of guanine residue in the opposite strand.

Apurinic site formation has also been detected, albeit indirectly, by high pressure liquid chromatography (HPLC) analysis as C8-aryl modified bases can be detected from solutions resulting from calf thymus DNA or C50 cells treated with with MBD, HMBD, or CBD[7,14]. As a result of C8-aryl purine adduct formation, the glycosidic bond is more vulnerable to hydrolysis and the depurination rate is increased. This suggested that aryl hydrazines may induce carcinogenesis via depurination of C8-aryl modified purine. This

depurination, facilitated by C8-aryl adduct formation, may explain the previous result that aryl hydrazines mediated DNA single strand breaks and DNA cross links in V79 cells. Nevertheless, further study needs to be conducted to explore this possibility.

Another possible mechanism that will be addressed throughout this dissertation is Z-DNA mediated aryl hydrazine carcinogenesis. As described in the previous section, several studies have suggested the biological role of Z-DNA in the transcription process. Furthermore, there is mounting evidence that indicates the mutagenic potential of Z-DNA. These results lead to the idea that Z-DNA "over" stabilization is likely to cause genetic alteration by disturbing transcription balance and/or increasing vulnerability of DNA to damage. Given the results from several studies that C8-guanine adducts can promote the conversion of B- to Z-DNA by destabilizing B-DNA and/or stabilizing Z-DNA, aryl hydrazine carcinogenesis through Z-DNA formation by C8-aryl adduct seems to be possible.

Previously[22] our group has shown that the C8-phenyl guanine adducts promote B-Z conversion in CG^{8Ph} ($d(5'\text{-}CGCGCG^{*}CGCG\text{-}3')_2$, G* = C8-phenyl modified guanine). The aryl modified DNA requires much lower NaCl concentrations to stabilize the Z conformation compared with the unmodified DNA, indicating the stabilization effect of C8-phenyl guanine adduct on Z-DNA. The result from this initial study confirmed the possibility that B-Z transition may play a role in aryl hydrazine carcinogenesis and has opened the opportunity of further explore the effect the C8-arylguanine adduct has on Z-DNA stabilization or B-DNA destabilization. Given that Z-DNA is suspected to play a significant role in cancer development, the C8-aryl adduct formed from the carcinogenic aryl hydrazines should, in general, promote the Z-DNA formation. To prove this concept

and to determine the correlation between Z-DNA formation and carcinogenesis, several C8-arylguanines, modified DNA including CG^{8Tol}, CG^{8HMPh}, CG^{8MMPh}, and CG^{8CPh} (d(5′-CGCGCG*CGCG-3′)$_2$, G* = 8-*p*-tolyl, 8-*p*-hydroxymethylphenyl, 8-*p*-methoxymethylphenyl, or 8-*p*-carboxyphenyl guanine) were made. These modified oligonucleotides are formed from the arenediazonium ions (MBD, HMBD, MMBD, and CBD, respectively), via aryl radicals, and all have been shown to be mutagenic[4,46]. In this study, the effects of each C8-aryl adduct on B/Z-DNA equilibrium was investigated.

The next chapter provides fundamental knowledge regarding methodologies and experimental techniques that have been used in this research. The experimental results are reported in Chapter 3 (DNA synthesis) and 4 (structural and conformational analysis). The computational study of the modified DNA, conducted to help explain the experimental results, is described in Chapter 5. Finally in Chapter 6, the effect of C8-arylguanine adducts on B/Z-DNA equilibrium and the significance of Z-DNA in aryl hydrazine carcinogenesis are discussed.

CHAPTER 2
EXPERIMENTAL TECHNIQUES

2.1 Synthesis of CG Decamers Containing C8-Aryl Modified Guanine

Generally, most short length oligonucleotide sequences can be efficiently made through use of an automated DNA synthesizer that utilizes phosphoramidite chemistry. The DNA synthesizer uses nucleoside cyanoethyl phosphoramidites (dA CE-PA, dG CE-PA, dT CE-PA, dC CE-PA, Figure 2.1) as the building blocks for the chemical synthesis of oligonucleotides. However, in order to make CG decamer with C8-arylguanine adduct, the modified C8-aryl-dG phosphoramidites (Figure 2.1) must be made as they are commercially unavailable. In this section, an overview of the chemistry behind automated DNA synthesis and the synthesis of C8-aryl-dG phosphoramidites will be described.

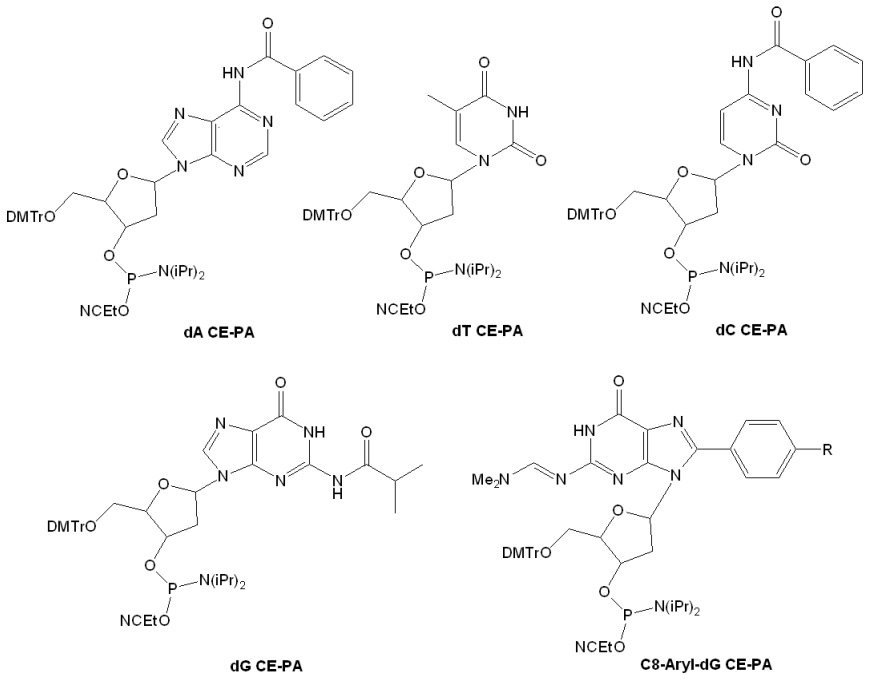
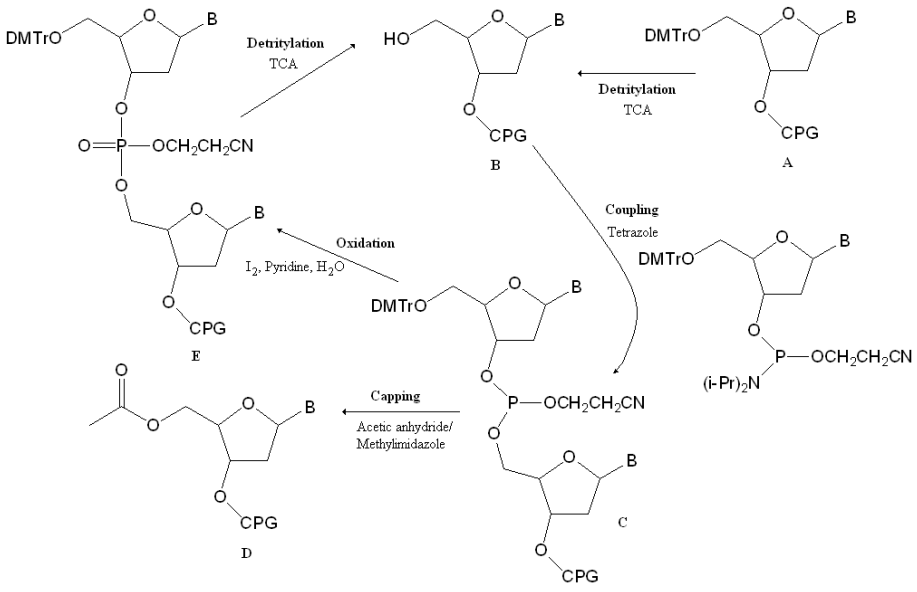


Figure 2.1 Structures of the phosphoramidites used in DNA synthesis.

2.1.1 Chemistry for Automated DNA synthesis[56]

Fundamentally, chemical synthesis of DNA works on the basis that the oligonucleotide can be elongated by coupling the reactive 3′-phosphoramidite of one nucleoside and the 5′-hydroxyl of the existing oligonucleotide. The synthesis occurs on the controlled pore glass (CPG) solid phase column with the first 3′ base attached to the column through an ester linkage[57]. All phosphoramidites, reactants, and solvents are flushed through the synthesis column with the oligonucleotide being extended from the 3′ end toward the 5′ end. Each base extension cycle is composed of four steps (Scheme 2.1) including detritylation of 5′-hydroxyl, base coupling, capping of unextended DNA, and oxidation. The reaction cycle repeats, for each base addition, until the DNA sequence is completed.



Scheme 2.1 The DNA synthesis cycle is including Detritylation, Coupling, Capping, and Oxidation. (B = A, G, C, T)

Detritylation deprotects the 5′-hydroxyl of the nucleoside bound to the solid phase and renders it a nucleophile for coupling in the next reaction (A→B). Trichloroacetic acid (TCA) in CH_2Cl_2, a protic acid, is used as a deprotecting agent as the DMTr group is acid labile. The protonation on the 5′-hydroxyl leads to DMTr cation loss and the detritylation can be driven to completion by removal of DMTr cations.

After removal of excess TCA in detritylation step, the phosphoramidite is activated by tetrazole which transfers a proton to the nitrogen of the diisopropyl group on the 3′-phosphorous making the protonated amine a better leaving group. The nucleophilic tetrazole then attacks the phosphorous to form a tetrazolyl phosphoramidite. The molar excess of tetrazole ensures that most of the phosphoramidite is activated. The solid supported end 5′-hydroxyl reacts with the 3′-tetrazolyl phosphoramidite to form the internucleotide phosphite linkage (B→C). As a result, the oligonucleotide is extended.

The coupling reaction between nucleotides may not proceed completely leaving some of the oligonucleotide attached to the solid support unextended. To prevent the unreacted oligonucleotide from being extended and giving rise to undesirable oligonucleotide sequences, unreacted material is capped. Acetic anhydride and methylimidazole are used as a capping reagent and are delivered to the synthesis column simultaneously. The two reagents react and give a reactive acetylating agent which, in turn, reacts with the 5′-hydroxyl of the unextended oligonucleotide to inactivate the moiety from further reaction (C→D).

The last step is to oxidize the freshly formed internucleotide trivalent phosphite to pentavalent phosphodiester. Iodine in an aqueous pyridine-tetrahydrofuran (THF) solution is used as a mild oxidizing agent. The iodine-pyridine complex forms a di-iodo

adduct with the phosphate that is subsequently hydrolyzed by water to form the internucleotide phosphodiester linkage.

To maximize DNA synthesis yield, each step of the synthesis cycle is followed by an acetonitrile (ACN) and argon flush to make sure that no reagents from the previous reaction remain to interfere with the next cycle. Following the oxidation of the final base, the DMTr protected 5′-hydroxyl group of the completed sequence can be detritylated. After the completion of the DNA synthesis the amine protecting groups and the solid support linkage are removed by incubating the synthesis column in a basic ammonium hydroxide solution for 15 hrs at 55°C. The crude oligonucleotide can be purified if necessary.

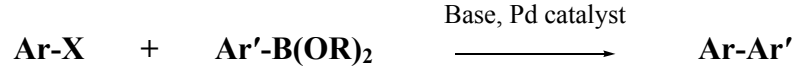
2.1.2 Synthesis of C8-Arylguanine Phosphoramidites

To incorporate C8-arylguanine adduct to the CG decamer, the modified C8-aryl dG phosphoramidites are required. The phosphoramidite used in the DNA synthesis contains several required functional groups including a diisopropylamino cyanoethyl phosphorous on 3′-hydroxyl, a DMTr on 5′-hydroxyl, and a protecting group of the exocyclic amine of nucleoside base. In case of the C8-aryl modified guanosine phosphoramidite, the purine ring has an additional aryl substituent at C8. Several reactions need to be performed in order to add these functional groups to prepare the aryl modified phosphoramidite. Our scheme starts with attachment of the aryl group to the C8 position of 2′-deoxyguanosine. Next the exocyclic amine is protected, followed by 5′-hydroxyl and 3′-hydroxyl groups are functionalized, respectively.

2.1.2.1 Synthesis of C8-Aryl dG Using the Suzuki Coupling

The 8-aryl-dG can be made through the chemical reaction between arenediazonium ion and dG as described by Kohda K. et al.[15] To make 8-phenyl-dG, the BD solution was freshly made from the reaction between aniline and $NaNO_2$. After reacted with BD in basic condition for 24 hrs, dG was arylated and 8-aryl-dG was collected from crystallization. Despite the success of making 8-phenyl-dG, using the carcinogenic arenediazonium ions as the arylating agents is quite hazardous requiring special handling. In addition, the reaction produces nitrogen which causes foaming, is difficult to contain, and to filter after the reaction is complete. It is also not very efficient, and provides relatively low yields (40%) in the case where the aryl group is phenyl. Finally, while this reaction works with some substituted phenyl groups it is unlikely to work with many substituted phenyl groups that are of interest. Improved methods that provide a safer, more convenient protocol, higher product yields and simple purification to isolate the product is preferred.

Palladium catalyzed cross coupling reactions have been successfully used for carbon modification of nucleosides[58,59]. Among the cross coupling reactions, Suzuki coupling between halonucleosides and arylboronic acids have been shown to be an effective way to make aryl modified nucleosides[60,61]. In general, the Suzuki coupling is the important carbon-carbon bond forming reaction that can be used to make biaryl compounds through palladium catalysis of the reaction between arylboronic acids or esters and aryl halides, sulfonates, or diazonium salts in basic condition[62,63] as shown in Scheme 2.2.



X = I, Br, Cl, OTf, OTs, OMs, $N_2^+BF_4^-$

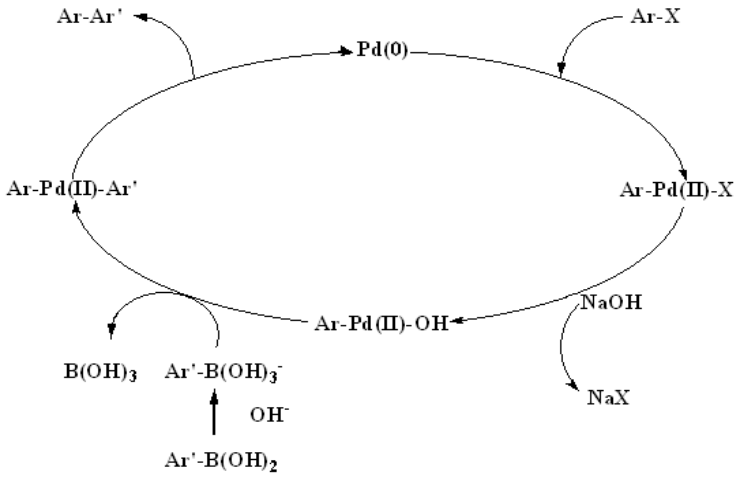
R = H or Alkyls

Scheme 2.2 The synthesis of biaryls from arylboronic acids or esters and electrophilic aryl reactants by Suzuki coupling

The general mechanism[64-66] of the coupling reaction is described in Scheme 2.3. The reaction begins with the arylhalide (or other electrophilic aryl reactant) reacting with the Pd(0) catalyst by oxidative addition to form an arylpalladium (II) halide. The hydroxide (or other negatively charge bases such as alkoxide, carbonate, or phosphate etc.) then attacks the boron atom in the arylboronic acid to give an aryl borate- *ate* complex. In addition, the nucleophilic base also reacts with the arylpalladium (II) halide to generate an arylpalladium (II) hydroxide. The aryl borate and the arylpalladium (II) hydroxide then react and form a diarylpalladium (II) complex which eventually undergoes a reduction-elimination reaction that leads to the formation of a biaryl product and the regeneration of the Pd(0) catalyst.

Several factors affect the efficiency of the Suzuki coupling reaction[62]. Because the reaction requires basic conditions, base selection is very important. Using different bases can give different products from the coupling. Weaker bases such as triethylamine (TEA) usually are less efficient with respect to product formation. The solvent system is also crucial since it strongly determines the solubility and basicity of the base that used in the coupling. In our case, making the modified nucleoside, that contains the hydrophilic sugar moiety, requires an aqueous solvent system. This limits the range of bases that can

be used to water soluble base. Generally, carbonate or hydroxide have been used and seem to work well.



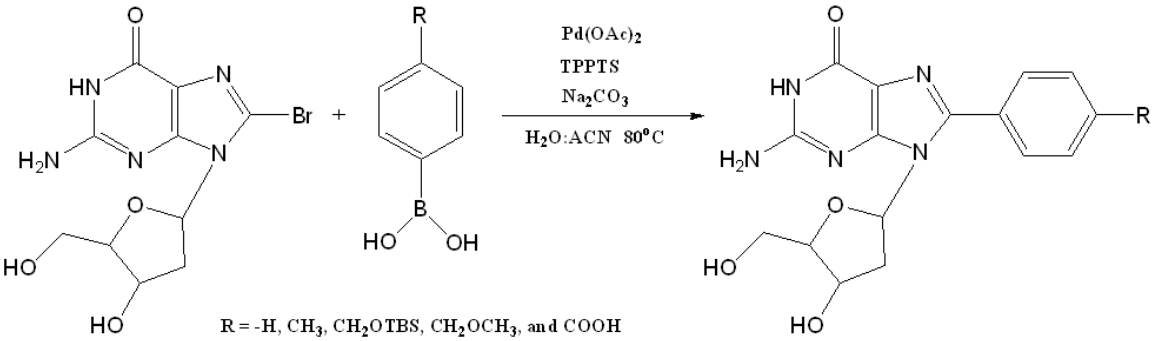
Scheme 2.3 The general mechanism of the Suzuki coupling reaction

The homogeneous mixture from water and water-miscible organic solvent such as ACN or THF can be used to improve the solubility of the hydrophobic reactants. In case this is necessary, the biphasic solvent mixture from organic solvent and water can be used if the lipophilic part is needed. The palladium catalyst is probably the most important part in the Suzuki coupling reaction. Pd(0) is the actual catalytic species but is usually generated *in situ* by reduction of Pd(II). Once formed, the Pd(0) must be kept in solution and, since it is sensitive to oxidation, the reaction is conducted under an inert gas such as argon or nitrogen to prevent reaction with oxygen. In addition, using a suitable and soluble palladium ligand is also essential to solubilize and stabilize Pd(0) in solution.

Shaughnessy et al.[67] have developed an the efficient way to synthesize biaryl compounds from an aqueous based Suzuki coupling using ACN as a co-solvent. The key discovery was that sterically demanding water soluble ligands can be successfully used

with palladium acetate ($Pd(OAc)_2$) to catalyze the coupling reaction. The idea seemed to be suitable for preparation of aryl modified nucleosides and has since been used to synthesize C8-arylpurine, as was demonstrated by Western E.C, et al.[61] in 2003. The successful aqueous Suzuki coupling of 8-Bromo-2′-deoxyguanosine (8-BrdG) and several arylboronic acids were reported in this study with the product yield in range of 70-90 %. The reaction is conducted at 80°C with Na_2CO_3 as the base in deoxygenated aqueous acetonitrile with $Pd(OAc)_2$ as the catalyst in a presence of the palladium ligand TPPTS (tris-(3-sulfonatophenyl)phosphine) ligand (Figure 2.4).

These studies provided a relatively efficient and convenient method to synthesize 8-aryl dG for the C8-arylguanine modified DNA synthesis. The protocols of making 8-aryl dG used in this study were followed and/or further developed.



Scheme 2.4 The synthesis of C8-aryl dG using the Suzuki coupling

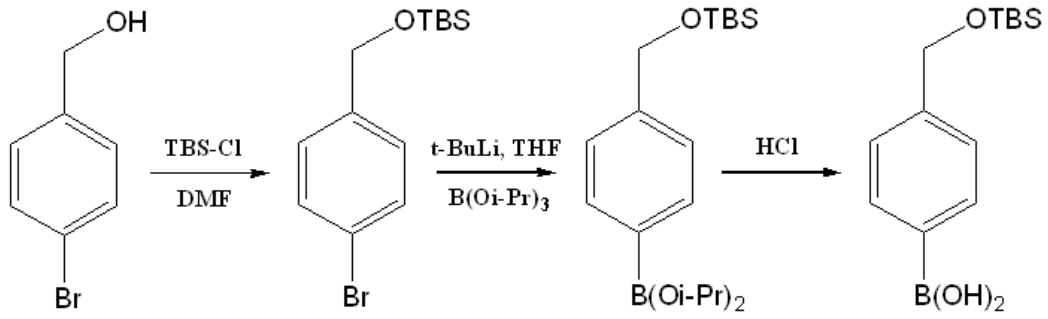
2.1.2.2 Synthesis of the TBS Protected *p*-Hydroxymethylphenylboronic acid

Most of the arylboronic acids needed here, including phenyl, *p*-tolyl, *p*-hydroxymethylphenyl, *p*-methoxymethylphenyl, and *p*-carboxyphenyl boronic acids, are commercially available and have been successfully used directly to make the modified

DNAs except for the *p*-hydroxymethylphenyl derivative. The 8-(*p*-hydroxymethylphenyl)-dG phosphoramidite has been synthesized, however using it in for the modified DNA synthesis failed to unequivocally produce the desired DNA, as suggested by mass spectral analysis. This synthesis failure may occur due to the nucleophilic nature of the benzyl hydroxyl group that can interfere with the DNA synthesis. In order to prevent the hydroxyl substituent from interfering with the DNA synthesis, a protecting group was needed on the hydroxyl group. The t-butyldimethylsilyl (TBS) protecting group has been used as a 2′-hydroxyl protection group of ribonucleosides[68] and is cleaved by exposure to concentrated ammonium hydroxide after DNA synthesis along with other protecting groups. Therefore TBS was chosen for use as the protecting group for the *p*-hydroxyl substituent.

The arylboronic acids can be made through the classical Grignard reaction[62] between aryllithiums and trialkyl borates. The reaction generates arylboronic esters which are hydrolyzed by acid to form the arylboronic acids[69]. From this reaction, one can vary the *p*-substituent of aryllithiums to make various *p*-substitute arylboronic acids. Zheng N. et al.[70] has shown the efficient way to make the TBS protected *p*-hydroxymethylphenylboronic acid in a pilot scale by utilizing a Grignard reaction. The synthesis protocol is shown in Scheme 2.5 starting with the protection of the hydroxyl of 4-bromobenzyl alcohol by using TBS-Cl to give the TBS protected product. The protected *p*-hydroxymethylphenylbromide is then reacted with t-buthyl lithium to form the TBS protected *p*-hydroxymethylphenyl lithium which is then treated with tri-isopropylborate to form the TBS protected *p*-hydroxymethylphenylboronic ester. After acid hydrolysis, the TBS protected *p*-hydroxymethylphenylboronic acid was isolated.

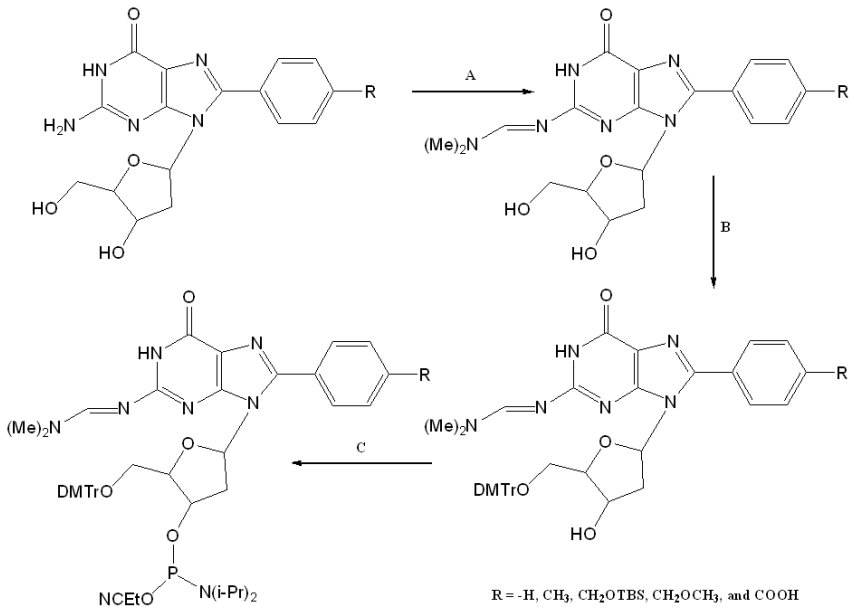
This final product was then used to prepare the modified guanine phosphoramidite which was successfully used to synthesize the hydroxymethylphenyl guanine adduct containing DNA. This arylboronic acid synthesis protocol is potentially useful for making other substituted arylboronic acids that may require protecting groups.



Scheme 2.5 The synthesis of 4-(TBS-O-methyl)-phenylboronic acid

2.1.2.3 Syntheses of Protected C8-Aryl dG and Phosphoramidites

Methods to prepare the C8 modified dG phosphoramidites for the oligonucleotide synthesis[15,71,72] have been published. The routes typically begin with C8-arylation of dG and are then followed by several reactions to prepare the modified dG phosphoramidite monomers for use in automated DNA synthesis as shown in Scheme 2.5.



Scheme 2.6 The synthesis of the phosphoramidite monomer. The reaction conditions are including A) N'N-Dimethylformamide dimethyl acetal, MeOH; B) DMTr-Cl, pyridine, TEA; C) 2-Cyanoethydiisopropylchlorophosphoramidite, CH_2Cl_2, TEA.

Various amine protecting groups including benzoyl, isobutyryl, and N,N-dimethylformamidine can be used to protect the N2 position of the aryl modified dG. The latter group has been shown to improve the stability of the nucleosides[73] toward acid and was used here for the synthesis of the C8-aryl dG phosphoramidites. To attach the N2-(N,N-dimethylformamidine) protecting group, N,N-dimethylformamide dimethyl acetal is reacted with the 8-aryl dG in dry methanol. The protecting group, which is a protected amide, forms a formamidine with the N2 amino group of the 8-aryl dG. The reaction has been found to be very sensitive to moisture. Water will hydrolyze the protected amide to N,N-dimethylformamide (DMF) which is inactive for the amine protection reaction. Generally, the methanol that used in the reaction must be dried over molecular sieves under inert gases such as nitrogen or argon. The reaction proceeds to the completion within 24 hrs, at room temperature, with product yields of approximately 90%. The

amine protected 8-aryl dG may be purified using normal phase column chromatography. However, the purification is usually unnecessary because the crude reaction product is quite pure as found by NMR of the crude reaction product. Simply drying *in vacuo* has proven to be sufficient to prepare the crude product for the next reaction.

After the protection of the N2 amine group, the 5′-hydroxy group of the 8-aryl dG is protected using DMTr-Cl, as is the case for unmodified phosphoramidites used in automated DNA synthesis. The reaction mechanism involves the nucleophilic attack (S_N1) of the 5′-hydroxyl group to the electrophilic tertiary carbon center of DMTr-Cl, resulting in the formation of the DMTr protected 5′-hydroxyl nucleoside and HCl. Since the nucleosides are sensitive to acid hydrolysis, the reaction is conducted in pyridine, a basic solvent, and in a presence of TEA to neutralize the acid generated and maximize the product yield. The reaction can be monitored by normal phase thin layer chromatography (TLC) with visualization by phosphomolybdic acid. The reaction product appears as an orange spot, likely due to cleavage of the DMTr protecting group from the product as the DMTr group is very acid labile and gives DMTr-OH which is orange in color. Unlike the previous protection step, the DMTr protection product requires purification. Several chromatographic systems have been explored for purification of the DMTr protected nucleosides[73]. A silica gel stationary phase was first used and was shown to cause decomposition of the nucleoside due to the acidic nature of the stationary phase. In contrast, Alumina, a basic stationary phase, has been found to be a more suitable system and causes minimal product decomposition. A mixture of CH_2Cl_2 and methanol gradient (0-10%) is used as the mobile phase and, after chromatography, the product yield is usually in range between 40% and 60%.

The attachment of a 2-cyanoethoxydiisopropylaminophosphoramidite at the 3′-hydroxyl position of the 8-aryl dG is the final step of the phosphoramidite synthesis. 2-Cyanoethyldiisopropylchlorophosphoramidite is used as the phosphatidylating agent. The reaction also involves the nucleophilic substitution of the chlorine on phosphorous by a 3′-hydroxyl of the 8-aryl dG. The reaction is conducted in CH_2Cl_2 in a presence of TEA to neutralize HCl that is generated during the course of the reaction. Due to high reactivity of the phosphatidylating agent, most nucleophiles, including water, can interfere with the reaction. Therefore, using dried, purified starting materials and solvent is necessary to achieve the desired product and optimize the yield, as is conducting the reaction under an inert atmosphere. The phosphoramidite is only partially purified by the precipitation of TEA hydrochloride salts out of the reaction by the addition of a mixture of benzene and THF. While further purification can be done by column chromatography, this step has been found to be unnecessary and tends to lead to the product degradation. After precipitation of salts, the solvents are removed *in vacuo* and the product dried over phosphorus pentoxide *in vacuo*.

By following these synthesis protocols, the aryl adducts with different *p*-substituents on phenyl ring can be introduced to the C8 position of dG. The 8-aryl dG phosphoramidites then were made and used for the modified CG decamer synthesis through the DNA synthesizer.

2.2 Purification of CG Decamers Containing C8-Aryl Modified Guanine

A number of factors, including low quality of the starting materials and reagents, degradation of the modified DNA during synthesis, machine malfunction, and other

technical issues, can affect the efficiency of the modified oligonucleotide synthesis. Although these unfavorable factors can be prevented or minimized, most of time they can not be completely eliminated. This will decrease the efficiency of the synthesis cycle and resulting in by-product generation. In order to get the sufficiently pure DNA that is pure enough for NMR and CD experiments, the crude modified oligonucleotides require purification.

Several HPLC techniques including ion-exchange chromatography and reversed phase chromatography[74] can be utilized to purify the modified oligonucleotides. The fast protein liquid chromatography (FPLC) method is one of the HPLC methods that has been widely used to isolate proteins in biochemistry and enzymology research and has also been used to purify the synthesized DNA. Because of the anion exchange matrix stationary phase that used in the preparation column, FPLC separates DNA mainly by the differences in negative charge of the DNA molecules. The oligonucleotide structures contain the phosphate backbones which carry negatively charges. Therefore, DNAs of various lengths will contain a different number of negatively charges and will bind to the anion exchange resin to a different degree. The longer DNA sequences that have more negative charge adsorb to the cationic stationary phase more strongly and require a more anionic mobile phase to elute them out of the column than shorter DNAs. In addition to the charge separation, there is some dependence of elution time on the hydrophopic nature of the DNAs with more hydrophobic sequences eluting more quickly.

The anion exchange matrix, diethylaminoethyl (DEME) modified resin, is commonly used to isolate nucleic acids[74] and have been successfully used to purify the modified CG decamer in this study. The quarternary ammonium ion functional groups on

DEME give them the anion exchanging capability. The counter ions, often OH^-, maintain electroneutrality of the stationary matrices. Once the oligonucleotides are loaded on to the column, the multi-negatively charged molecules will replace the previous counter ions with stronger electrostatic forces. A mobile phase that contains anions, in this case Cl^-, has been used extensively to elute the oligonucleotides and a gradient comprised of sodium chloride and sodium hydroxide can effectively be used to separate the modified oligonucleotides contained in the crude DNA synthesis mixture. Generally, the desired product, the longest sequence in the synthesis process, is the last one that is eluted off of the column since it bares the highest negative charge. The separation of the crude DNA can be monitored by using an ultraviolet (UV) light detector at 260 nm where nucleic acids generally have the highest absorbtivities.

After isolation by FPLC, the oligonucleotides still need to be isolated from the sodium chloride and sodium hydroxide salts that were present in the eluant. Solid phase extraction (SPE) using reversed phase C18 cartridge can be used to desalt the DNA samples. Because of the partial hydrophobic parts of the DNA (the purine and pyrimidine rings) the oligonucleotides can bind to the lipophilic hydrocarbon chain in the stationary phase better than NaCl and NaOH. Therefore, the polar impurities can be eliminated by washing the cartridge with water after the DNA samples were loaded. The DNA can be eluted with an aqueous methanol mixture. After removing the solvent, the purity of the modified oligonucleotides can be confirmed by FPLC analysis. A single peak in the FPLC chromatogram is expected if the synthesized oligonucleotides are pure.

2.3 Mass Spectrometry of CG Decamers Containing C8-Aryl Modified Guanine

Mass Spectrometry (MS) analysis has been commonly used to characterize organic compounds and study molecular interaction of a wide variety of chemicals. This method of analysis is based on the detection of weight to charge ratio of the ionized molecules by measuring the response of their trajectories to electronic or/and magnetic fields. Hence, in the process of detection, the analytes need to be ionized *in vacuo* and move through an electronic field to the mass analyzer which separates the ionized molecules by their mass to charge ratio (m/z). The detector will then detect the sorted ionized molecules and convert the signal into a mass spectrum that is generally distinct for each chemical.

The classical MS methods usually use electron (EI) or chemical ionization (CI) as an ionization source[75]. The high energy electron beam used in EI to interact with the analytes and generate the molecular ions ($M^{+}\cdot$) and daughter ions. Unlike EI-MS in which electrons directly interact with the analyte molecules, CI-MS method uses an ionized reagent gas such as protonated methane (CH_5^+), isobutene cation ($(CH_3)_3C^+$), or ammonium ion (NH_4^+) as the ionization reagent to generate protonated analytes (MH^+). However, due to the limitations of conventional techniques, it is difficult to convert thermal labile and/or non-volatile macro molecules such as nucleic acids or proteins into the ionized gaseous state without decomposition.

The development of soft ionization methods including electrospray ionization (ESI) and desorption ionization have provided the opportunity to use MS as an analytical tool for biological molecular research. ESI-MS analysis has been successfully used in for the characterization and sequencing of biological polymers and can be used to probe

molecular interaction[76,77]. Here, we have used ESI-MS to characterize the C8-aryl dG monomers and the C8-arylguanine modified CG decamers.

The ESI-MS method produces the quasi-molecular ions[78] by nebulizing a sample solution, in which analytes are already in the ionic form, into an aerosol through a high voltage chamber. The charged droplets containing the analytes that are generated are subjected to the counterflow of a drying gas, generally nitrogen gas. This process helps to accelerate solvent evaporation, decreases the size of the droplets, and increases the charge density on the droplet surfaces which ultimately causes the droplets to burst or undergo a "Coulomb explosion" due to their electrostatic repulsion. As a result, daughter droplets are produced and further break down until the preformed ions are generated (Figure 2.2). The preformed ions[79] are the ionic analyte droplets that are small enough (ideally each droplet contains only one molecule of analyte) so they can be desorbed into the gas phase and pass through the capillary tube into the m/z analyzer. The ion adducts can be incorporated into the preformed ions in cases where the analyte molecules are not well dissociated or neutral. For instance, ammonium acetate buffer can be used as a source of the ammonium ion adducts which give positive charges to the preformed droplets.

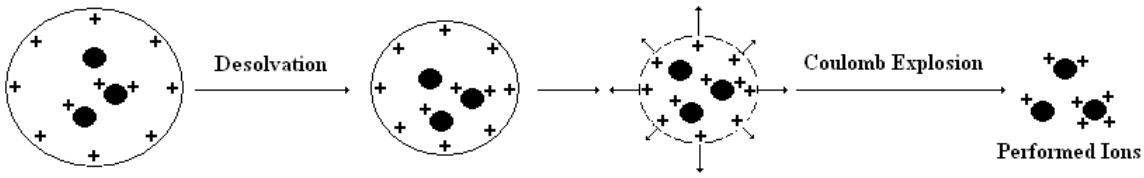


Figure 2.2 A diagram shows the ionization process in the positive ion mode ESI-MS to generate the preformed ions

Depending on the mode of detection, ESI-MS can be used to detect either negatively charged ions in a negative ion mode or positively charged ions in a positive ion mode. Fundamentally, the ionization state of the analyte molecule is mainly dependent on the solvent system and the pKa of the analyte molecule. Volatile aqueous organic solvent (methanol, ACN, isopropylalcohol, etc.) systems are commonly used in ESI-MS. In addition, a limited amount of volatile acids or bases, such as formic acid, acetic acid, trifluoroacetic acid, ammonium formate, ammonium acetate, ammonium hydroxide, and TEA, may be used as buffers to adjust the optimum pH of the solvent system. Non-volatile solvents or salts should be avoided since they may accumulate in the system and cause machine malfunction. In general, acidic molecules from negative ions in high pH solution while the basic molecules would be protonated and form positive ions in low pH solution. If the analyte is in a neutral form at the working pH, the ion adducts from buffer ions will largely determine the charge of the ions.

Multiple charge states of the ions are commonly observed during the ionization process and depend on the structure of the analyte and the solvent system. This phenomenon is actually advantageous for the analysis of the biological macromolecules like peptides or oligonucleotides that have high molecular weight (MW). Because ESI-MS detects m/z of the preformed ions, the higher charge state of the analyte the lower m/z ratio compared to a lower charge state of the same analyte. Unlike the conventional ionization methods, ESI can be used to analyze high MW compounds up to 100,000 mass units due to its ability to generate ions with multiple charged. The mass spectra of the biological molecules from ESI-MS are usually composed of a series of peaks from several charge states (M^{n-} or M^{n+}, n is a charge state of the molecule). The MW of the

analyte then can be determined from a mathematically approximation from the resulting mass spectrum.

In the positive ion mode, the molecular mass (M) of the analyte can be calculated[80] from a grouping of m/z ratios of the ion series and *n* as described by the following equations.

From	$m/z_1 = (M + n)/n$	Equation 2.1
and	$m/z_2 = (M + n + 1)/(n + 1)$	Equation 2.2
Therefore	$n = (m/z_2 - 1)/(m/z_1 - m/z_2)$	Equation 2.3

By solving Equation 2.1 after calculating *n* from Equation 2.3, the molecular mass of the analyte can be estimated. Similar to a positive ion mode, M can also be calculated in a negative ion mode from the observe m/z and *n* as described below in the Equations 2.4 to 2.6.

From	$m/z_1 = (M - n)/n$	Equation 2.4
and	$m/z_2 = (M - n - 1)/(n + 1)$	Equation 2.5
Therefore	$n = (m/z_2 + 1)/(m/z_1 - m/z_2)$	Equation 2.6

The m/z ratios can be obtained from the mass spectrum of the analyte and used to calculate *n* from the Equation 2.6. After replacing the calculated *n* in Equation 2.4, M of the analyte then can be resolved. The calculation may need to be justified accordingly in case that the ion adducts are present. Alternatively, the series of m/z of multiple charge

states of the analyte, with or without the ion adducts, can be calculated based on the expected MW of the analyte and compared to the observed m/z in order to characterize the sample.

The negative ion mode seems to be a suitable detection method for the oligonucleotides mass analysis because of the net negative charge of the DNA phosphodieater backbone (pKa ~ 1)[77] which are usually dissociated in a sample solution at pH 7.4. Hence, without any buffer, an aqueous methanol solvent system for the modified CG decamers analysis and the negative ion mode ESI-MS has been successfully used to characterize the oligonucleotides. A series of charge states of the oligonucleotides were observed and compared to the calculated value to verify the identity of the synthesized oligonucleotides.

2.4 Circular Dichroism Spectrophotometry of CG Decamers Containing C8-Aryl Modified Guanine

2.4.1. CD of B and Z DNA

CD spectroscopy is based on the fact that an optically active, asymmetrical molecule intereacts with the right and left handed circular polarized lights differently. When right and left circular polarized lights pass through absorbing optically active matter, not only do they travel at different speeds, they also absorbed to a dissimilar extent. The difference between the circular light absorptions ($\Delta\epsilon = \epsilon_L - \epsilon_R$) is called circular dichroism[81] which is detected in the CD spectrometer. When two circular polarized light beams, with equal amplitudes and wavelengths, are superposed the result is a linear polarized wave. In case of CD, the superposition of the right and left polarized

light is no longer a linearly polarized wave. Instead, elliptically polarized light is generated. The ellipticity that is caused by CD can be mathematically determined and is usually presented in a CD spectrum. CD spectroscopy has been widely used to study molecular conformation of the chiral compounds including peptides and nucleic acids in several research areas due to the fact that different conformations of the same optically active substance can be distinguished by the unique patterns of their CD spectra.

The structure of nucleic acids is composed of the aromatic purine/pyrimidine bases that are connected through the internucleotide sugar-phosphate backbone. The phosphate groups have high energy electronic transition at wavelengths shorter than 170 nm. Close to the absorption band of phosphates, deoxyribose sugars absorb light begin at 190 nm. The nucleotide bases, A, G, C, and T, are the chromophoric parts of DNA because of the aromatic system of purine and pyrimidine bases. Since the nucleotide base contains several π bonds, it is possible for them to have a large number of $\pi \rightarrow \pi^*$ transitions which result in a lower electronic transition energy begining at 300 nm[81]. While the nucleotide bases are chromophoric, they are also symmetrical (plane of symmetry) and are not optically active which, in turn, makes them CD inactive. The deoxyribose sugar that attach to the base is chiral and gives the nucleoside molecule asymmetry and CD activity. However, since the absorbing moiety (pyrimidine or purine) is attached to the chiral component (deoxyribose) the effect of a chiral sugar on the CD is greatly attenuated and therefore, the CD intensity of a nucleoside base is relatively low. Several interactions including base-paring through hydrogen bonding, base stacking, hydrophobic and electronic interactions give rise to a super asymmetric species, the double helical structure of DNA which has high CD intensity. Due to the polymorphic

nature of nucleic acids and numerous secondary structures of DNA, other than the conventional B-DNA structure, can be adopted under specific conditions including the conformation of interest here, Z-DNA. Due to the differences in interstrand and intrastrand base-base interaction, B- and Z-DNA each have a unique 'signature' in the CD spectrum that can be used to distinguish the two structures.

The B to Z transition of several oligonucleotides including unmodified and modified poly (dCdG) have been investigated by using CD spectrophotometry[28,82-84] due to the fact that the repeated CG sequences can, relatively easily, assume the Z conformation. In general the CD spectrum of B-DNA shows the positive ellipticity at approximately 280 nm and 220 nm and a negative ellipticity at approximately 250 nm (Figure 2.3). In contrast, the CD spectrum of Z-DNA shows a sharp negative ellipticity at approximately 295 nm[81]. Based on the different CD patterns of B and Z-DNA, the conversion of B to Z-DNA can be monitored as demonstrated by Pohl, F.M. and Jovin, T.M. in 1972[82]. The study has shown that the conformational change of poly (dCdG) from B to Z can be achieved by stabilizing the Z conformation by increasing sodium chloride concentration. The transition process is completely reversible since the Z-DNA converts back to B-DNA after salt removal. At physiological pH with low salt concentration, unmodified oligonucleotides preferentially adopt the B conformation. Z-DNA formation can be observed at high salt concentration (more than 2M sodium chloride for unmodified oligonucleotides). A reduction of the CD signal at approximately 295 nm (trough formation) and an increase of CD at approximately 270 nm can be used to follow the transition of B to Z-DNA and to quantitate the conversion.

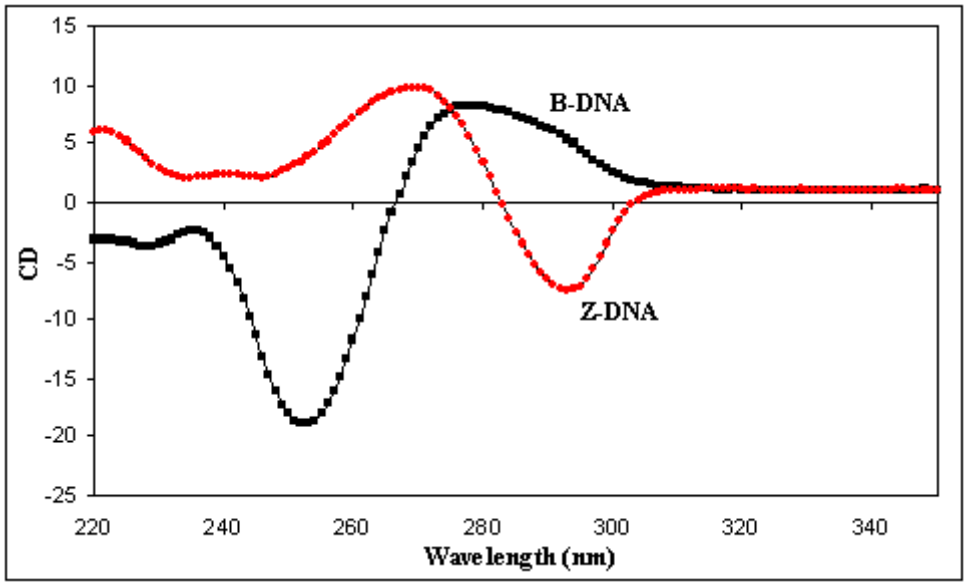


Figure 2.3 CD spectra of B-DNA (black) and Z-DNA (red) forms of the duplex CG decamer

To evaluate the effect of the C8-arylguanine adduct on B-Z transition, a series of CD spectra of the unmodified and aryl modified CG decamers in pH 7.4 phosphate buffer solution with various concentrations of sodium chloride have been collected. The B to Z conversion of the oligonucleotides can be observed from the transition of CD spectra as explained previously. The CD spectra can be used to quantitate the amount of the Z and B forms and thereby the effect of the C8-arylguanine adducts on Z-DNA stabilization and/or B-DNA destabilization can be roughly estimated in response to increasing sodium chloride concentration that needed for Z-DNA formation. The lower the concentration of sodium chloride that is required to cause the B to Z transition of the modified oligonucleotides, then, can be related to the higher stabilization effect of the C8-arylguanine adduct on a Z conformation or destabilization effect on B conformation of CG decamer. However, this method is only appropriate for preliminary comparison. A

more quantitative method that has been used to evaluate the effect of the aryl adduct on B-Z transition is discussed in the next section.

2.4.2. Z-DNA Quantitation

In order to compare the effect of the C8-arylguanine adduct on B- and Z-DNA stability in a more quantitative fashion, molar fractions of B, Z and ssDNA (f_B, f_Z, and f_{ss} respectively) can be calculated from the CD spectra. By quantitating the amount of Z-DNA in a sample, the Z-DNA stabilization effect of the aryl adducts can be inferred from the amount of Z-DNA that forms in the solution. The mathematical estimation of f_B, f_Z, and f_{ss} from CD spectra was reported previously by Xodo, L.E. et al.[85] and was used in this study. The calculation was based on the assumption that there is an equilibrium amongst three conformations of the oligonucleotide under study (B, Z, and ss). Therefore, at any given wavelength, the total molar ellipticity ($\Delta\epsilon$) of the oligonucleotide is equal to the sum of the molar ellipticities of each conformation, weighted by the molar fraction of each species (Equation 2.7). The CD at 295 nm has being selected here to use in the calculation, mainly due to the strong CD intensity of the Z conformation relative to the B form. The total molar fraction of all conformations equals the summation of f_B, f_Z, and f_{ss}, and, therefore, equals 1 (Equation 2.8).

At 295 nm $$\Delta\epsilon^{295} = \Delta\epsilon_B{}^{295} * f_B + \Delta\epsilon_Z{}^{295} * f_Z + \Delta\epsilon_{ss}{}^{295} * f_{ss} \quad \text{Equation 2.7}$$

$$1 = f_B + f_Z + f_{ss} \quad \text{Equation 2.8}$$

However, in practice, the instrument provides the total ellipticity (ϵ), not $\Delta\epsilon$. Therefore,

to make the calculation possible Equation 2.7 has been modified as shown in Equation 2.9 as shown below.

At 295 nm $CD^{295} = (\Delta\epsilon_{ss}^{295} * C_{ss}) + (\Delta\epsilon_{B}^{295} * C_{B}) + (\Delta\epsilon_{Z}^{295} * C_{Z})$ Equation 2.9

Where CD^{295} is the observed CD intensity at 295 nm and C_{ss}, C_B, and C_z are the concentrations of ssDNA, B-DNA, and Z-DNA in solution, respectively. To simplify the calculation, the path length (*b*) which is constant at 1 mm for all experiments is left out of the equation. The $\Delta\epsilon$ of each conformation can be estimated by assuming under selected conditions, most of the oligonucleotides can be driven into a single polymorphic form. At temperature as high as 90°C most of dsDNA will be denatured and is in the ssDNA form, thus the CD^{295} signal will be due to only the ssDNA ($CD^{295} = \Delta\epsilon_{ss}^{295} * C_{ss}$). Then, $\Delta\epsilon_{ss}^{295}$ can be solved with the observed CD^{295} and C_{ss} which can be quantitated by UV. With low salt concentration at room temperature, B-DNA is expected to be a sole conformation present and $\Delta\epsilon_{B}^{295}$ can be approximated from CD^{295} of the sample. By increasing a sodium chloride concentration to 2M-4M and cooling the sample to 10°C, the $\Delta\epsilon$ of Z-DNA can also be estimated the same way as done for the B and ssDNA. With the observed CD^{295} and the estimated molar ellipticities, C_{ss}, C_B, and C_Z are left to be solved. Note that C_{ss}, C_B, and C_Z are measurable by UV only when a single conformation exists in the sample under controlled conditions. Otherwise, they are considered as variables that need to be mathematically calculated. To get C_Z, both C_{ss} and C_B must be determined first.

The C_{ss} can be obtained by the following calculation. At the isosbestic point of

dsDNA, the molar absorptivities ($\Delta\varepsilon$) of B and Z-DNA are equal ($\Delta\varepsilon_B = \Delta\varepsilon_Z = \Delta\varepsilon_{ds}$).

Therefore $$\varepsilon^{270} = \varepsilon_{ds}{}^{270} * f_{ds} + \varepsilon_{ss}{}^{270} * f_{ss}$$ Equation 2.10

And $$1 = f_{ds} + f_{ss}$$ Equation 2.11

Similar to CD, the observed UV absorbance (A) has being used in the calculation instead of $\Delta\varepsilon$. Hence, Equation 2.10 has been modified to give Equation 2.12.

At 270 nm $$A^{270} = (\varepsilon_{ss}{}^{270} * C_{ss}) + (\varepsilon_{ds}{}^{270} * C_{ds})$$ Equation 2.12

Where A^{270} is the observed UV absorbance at 270 nm and C_{ds} is a concentration of dsDNA which equals to $C_B + C_Z$. The molar absorptivities of ss- and dsDNA can be obtained the same way as the molar ellipticities. At 90°C, $\varepsilon_{ss}{}^{270}$ can be estimated while $\varepsilon_{ds}{}^{270}$ is attained at low temperature. To further solve the calculation, the relationship between the concentrations of ssDNA and dsDNA in the sample is must be established. The equilibrium between ssDNA and dsDNA is varied based on a given condition, however the amount of total DNA is constant. Considering that at low temperature where DNA denaturation is minimized, we may assume the total concentration of oligonucleotide (C_{total}) as double helix based on the UV measurement.

Therefore $$C_{total} = (C_{ss}/2) + C_{ds}$$ Equation 2.13

And $$C_{ds} = C_{total} - (C_{ss}/2)$$ Equation 2.14

By replacing C_{ds} in Equation 2.12 with $C_{total} - (C_{ss}/2)$ (Equation 2.14) C_{ss} can be solved from Equation 2.15 (derived below).

$$A^{270} = (\varepsilon_{ss}{}^{270} * C_{ss}) + (\varepsilon_{ds}{}^{270} * (C_{total} - (C_{ss}/2)))$$

$$A^{270} = (\varepsilon_{ss}{}^{270} * C_{ss}) + (\varepsilon_{ds}{}^{270} * C_{total}) - ((\varepsilon_{ds}{}^{270}/2) * C_{ss})$$

$$A^{270} = ((\varepsilon_{ss}{}^{270} - (\varepsilon_{ds}{}^{270}/2)) * C_{ss}) + (\varepsilon_{ds}{}^{270} * C_{total})$$

$$A^{270} - (\varepsilon_{ds}{}^{270} * C_{total}) = (\varepsilon_{ss}{}^{270} - (\varepsilon_{ds}{}^{270}/2)) * C_{ss}$$

$$C_{ss} = \frac{A^{270} - (\varepsilon_{ds}{}^{270} * C_{total})}{\varepsilon_{ss}{}^{270} - (\varepsilon_{ds}{}^{270}/2)} \qquad \text{Equation 2.15}$$

Equation 2.9 can then be simplified to solve C_B by replacing C_Z with C_{ds}-C_B (because $C_{ds} = C_Z + C_B$). As a result, C_B can be calculated from Equation 2.16 which is derived as follows.

$$CD^{295} = (\Delta\epsilon_{ss}{}^{295} * C_{ss}) + (\Delta\epsilon_B{}^{295} * C_B) + (\Delta\epsilon_Z{}^{295} * (C_{ds} - C_B))$$

$$CD^{295} = (\Delta\epsilon_{ss}{}^{295} * C_{ss}) + (\Delta\epsilon_B{}^{295} * C_B) + (\Delta\epsilon_Z{}^{295} * C_{ds}) - (\Delta\epsilon_Z{}^{295} * C_B)$$

$$CD^{295} = (\Delta\epsilon_{ss}{}^{295} * C_{ss}) + (\Delta\epsilon_Z{}^{295} * C_{ds}) + (\Delta\epsilon_B{}^{295} * C_B) - (\Delta\epsilon_Z{}^{295} * C_B)$$

$$CD^{295} = (\Delta\epsilon_{ss}{}^{295} * C_{ss}) + (\Delta\epsilon_Z{}^{295} * C_{ds}) + ((\Delta\epsilon_B{}^{295} - \Delta\epsilon_Z{}^{295}) * C_B)$$

$$C_B = \frac{CD^{295} - (\Delta\epsilon_{ss}{}^{295} * C_{ss}) - (\Delta\epsilon_Z{}^{295} * C_{ds})}{\Delta\epsilon_B{}^{295} - \Delta\epsilon_Z{}^{295}} \qquad \text{Equation 2.16}$$

After C_B is solved, C_Z can then be estimated, and f_B, f_Z, ,and f_{ss} can be obtained accordingly.

The molar fractions of the C8-arylguanine modified oligonucleotides in all three conformations have being calculated based on the described method. Measurements were

made on DNA samples with a wide range of salt concentrations and temperatures. The resulting data from the CD experiments were used to generate a series of conformational diagrams which are useful to evaluate and compare the Z-DNA stabilization and B-DNA destabilization effects among various C8-arylguanine adducts. Then, comparing under the same set of conditions, the molar fractions of Z form of the different synthetic modified oligonucleotides can be used to compare the Z-DNA stabilizing effect (or may be B-DNA destabilizing effect) of the DNA adducts, given that more stable Z-DNA would have higher f_Z which indicates a greater stabilization effect from the C8-arylguanine adduct. CD spectroscopy has proven to be a useful technique to study the conformation and stability of the modified oligonucleotides.

2.5 Nuclear Magnetic Resonance Spectrometry of CG Decamers Containing C8-Aryl Modified Guanine

Nuclear magnetic resonance (NMR) is one the most important tools available for obtaining structural information of molecules. NMR has been used in this study to characterize synthetic nucleosides and also to determine the conformation of unmodified and modified oligonucleotides. Although CD can provide information about the overall conformation of an oligonucleotide, it cannot provide information regarding the local conformation. To confirm the results from CD experiments and to obtain a better understanding of the B-Z interconversion, several NMR experiments have been conducted. The principles of the NMR experiment will be discussed briefly in this section along with the NMR methods that have been used to study the B-Z equilibrium of the modified oligonucleotides.

2.5.1. Basic Theory of NMR[86-88]

All nuclei contain electrons that have the property of spin which may generate a magnetic dipole along the nuclear axis. Fundamentally, NMR spectra of the nuclei that have a non-zero spin number (I) , can be measured , especially nuclei that have I = ½ (e.g. ^{1}H and ^{13}C) due to their uniform spherical charge distribution. Based on quantum mechanics, the spin number determines the number of possible orientations ($2I$ +1) of nuclei when an external magnetic field is applied. In the case of a nucleus with I=½, there are two energy levels possible as shown in Figure 2.4.

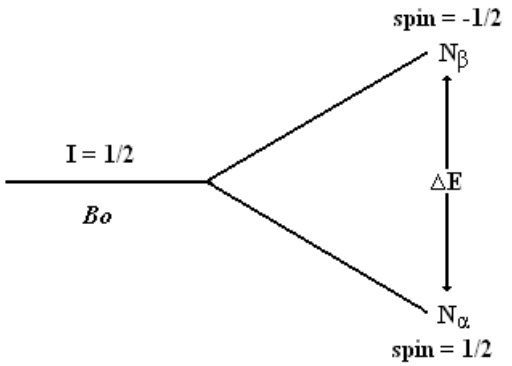


Figure 2.4 Two energy levels of the nuclei with I = 1/2 in a magnetic field strength of B_0

According to the Boltzmann distribution, the population of the lower energy state nuclei (N_α) is slightly greater than the higher energy state population (N_β). The energy gap (ΔE) between the two states is determined by a magnetic field strength (B_0) as shown in Equation 2.17,

$$\Delta E = (h\gamma/2\pi)B_0 \qquad \text{Equation 2.17}$$

where h is Plank's constant and γ is a gyromagnetic ratio which is constant for each atomic type of nuclei. In a static magnetic field, the populations of two energy states are established and the application of RF can be used to cause a transition between the two states. Given $\Delta E = hv$, the relationship between the applied radio frequency (v) and B_0 can be established as described in the Equation 2.18.

$$v = (\gamma/2\pi)B_0 \qquad \text{Equation 2.18}$$

By applying a matching RF at a particular magnetic field strength, nuclei in the lower energy state will absorb energy and rise to a higher energy level. The excited nuclei then relax back to the original state by radiating the absorbed energy which can be detected ultimately and recorded as the NMR spectrum.

The potential of NMR for structure determination was not recognized until the chemical shift effect was discovered in 1953[87]. The chemical shift effect arising from the electron cloud associated with nuclei plays a very important role in the observed NMR behaviour of other nearby nuclei. The electron spin of each nucleus generates its own local magnetic field that alters the applied magnetic field (B_0). Due to the shielding effect of the nuclei's electron cloud that alters the local B_0, the effective frequency (v_{eff}) is decreased as shown in Equation 2.19. A shielding constant (σ) is proportional to the degree of shielding by a nuclei electron cloud. The higher v_{eff} means less shielding effect from the electron cloud and the lower v_{eff} indicates more electron shielding effect.

$$v_{eff} = (\gamma/2\pi)B_0(1-\sigma) \qquad \text{Equation 2.19}$$

Because the chemical environment can affect the electron density of nuclei, identical nuclei that are in different environments will have distinctive shielding constants that can be recorded by NMR instrumentation and used to differentiate them. As a result, a series of absorption peaks, which represent nuclei in different environments, are observed in the NMR spectrum. The chemical shifts that are assigned in NMR spectrum indicate the absorption position of the nuclei in dissimilar environments compared to the reference nuclei. For instance, a hydroxyl proton will have a different chemical shift as compared to the protons from tetramethylsilane (TMS). The unit of chemical shift is in Hz which can be converted to the ppm unit, typically used for reporting the NMR spectrum. It is derived by dividing the observed frequency of nuclei by the operating frequency (MHz range) of the NMR spectrometer. Usually, a simple one dimension (1D) NMR spectrum shows a plot of the chemical shifts in ppm against the intensity of the absorption peaks of different nuclei. Furthermore, the peak area represents the relative number of each type of nuclei in a molecule.

In addition to chemical shift, spin coupling is an important phenomenon that provides additional information regarding chemical structure. The coupling constant (*J*) and multiplicity of the absorption peaks results from spin-spin coupling that can be used to assist chemical shift assignment in order to characterize a molecular structure of chemical, especially in ^{1}H NMR. An absorption peak of a certain proton can be affected by its magnetically active neighboring nuclei that typically are separated by two to three bonds. In a magnetic field, the energy states of nuclei are established based on their spin number. An NMR active nucleus can interact with nearby nuclei resulting in a multiplicity of absorption peaks for the nucleus under consideration. For example, as

shown in Figure 2.5, a triplet peak is observed for a proton that has two equivalent neighboring protons since the neighboring nuclei have three energy states that produce three different microfields. The coupling constant, which can be measured from the multiplet, is the separation between the two coupled NMR peaks and can be used to determine the relationship between protons (e.g., dihedral angle).

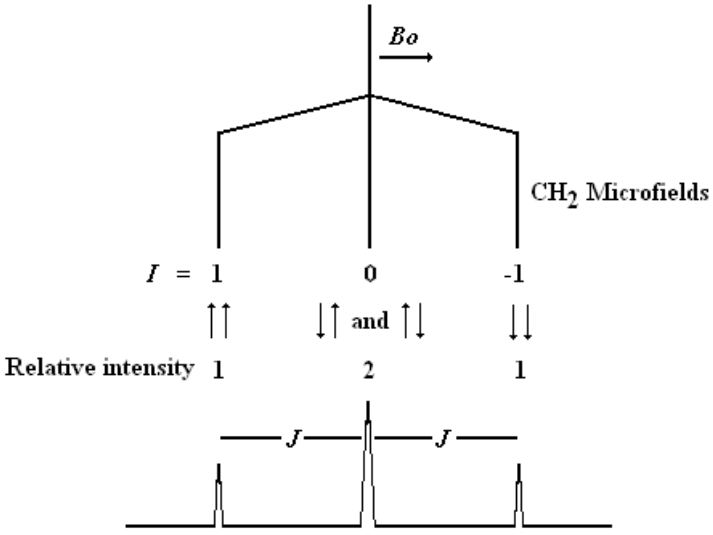


Figure 2.5 Triplet generated by neighbor CH_2 microfields

A simple first order multiplicity can be used to predict the pattern of proton absorption peaks based on the formula $n + 1$, where n is the number of adjacent protons with the same coupling constant. A more general formula, $2nI + 1$, can be used to cover all nuclei (i.e., spin > ½). However, if the coupling is non-first order (non-first order coupling is observed when then $\Delta\nu/J < 10$) more complicated coupling patterns may be observed. Overall, by gathering all of the pieces of information available from the NMR spectrum of a compound, including the chemical shifts, coupling constants, and multiplicities of

the absorption peaks, one may use the observed ^{1}H NMR spectrum (and perhaps additional spectra obtained on on nuclei other than protons) to assign both structure and conformation.

Because many organic compounds are comprised mainly of a carbon skeleton, ^{13}C NMR spectrometry has also been used to characterize the organic chemicals. The basic concept of ^{13}C NMR is the same as ^{1}H NMR. In practice, however, since the spin couplings between ^{13}C nuclei are very weak due to the low natural abundance of ^{13}C (1.11%), and therefore the likelihood of having two adjacent ^{13}C atoms, necessary to observe coupling, is low (^{12}C (~98.89%) is NMR silent). Furthermore, ^{13}C are typically obtained with proton decoupling and ^{13}C NMR spectra appear as a series of ^{13}C absorption singlets.

Generally, simple 1D spectra of ^{1}H and ^{13}C NMR may be sufficient to identify a simple organic molecule. However, in the case of a more complicated molecule, additional NMR experiments may be necessary. Two dimensional (2D) NMR provides enhanced capability to elucidate the structure of organic compounds. The correlation NMR experiments including COSY (**co**rrelation **s**pectroscop**y**) and HETCOR (**het**eronuclear **cor**relation spectroscopy) have been used extensively in this study to characterize the synthesized nucleosides. In addition to COSY, NOESY (**n**uclear **O**verhauser **e**ffect correlation **s**pectroscop**y**) has been used to study and assign the structure of the modified CG decamers and the B-Z DNA equilibrium.

A proton-proton coupling correlation, ^{1}H-^{1}H COSY, is based on spin couplings among protons and therefore reveals connectivity. The diagonal of the correlated spectrum corresponds to the normal ^{1}H NMR spectrum (typically displayed along both

the vertical and horizontal axes). The off-diagonal elements of a COSY spectrum indicate the couplings between protons. Since the magnitude of coupling constants falls off rapidly, the appearance of COSY correlations typically indicates the protons are adjacent. Similar to ^{1}H-^{1}H COSY, the ^{1}H $^{-13}$C COSY or HETCOR shows the correlation between ^{13}C and ^{1}H spectra. The cross peaks indicate the couplings between carbons and protons (and therefore attached protons) that may help identify unknown carbons or protons that can not be assigned from normal 1D NMR or assigned only with difficulty. With COSY and HETCOR, most protons and carbons of the nucleosides can be assigned.

Unlike COSY, 2D ^{1}H-^{1}H NOESY experiment utilizes the nuclear Overhauser enhancement (nOe) to obtain spectra that are correlated based on a through space interaction. To explain nOe, we consider two nuclei that are close in space and involved in the relaxation process of each other. If a saturating RF were applied to equalize the population of the two energy levels ($N_\alpha = N_\beta$) of one nuclei, the energy state dynamics of another nuclei will be affected resulting in the enhancement of integrated signal of the nuclei[89]. The complete mathematical explanation of nOe has been described by Solomon, I, 1955[90]. The nOe effect can be observed in 1D ^{13}C NMR. As proton decoupling is typically used, the absorption peak intensity of ^{13}C can be increased significantly (up to 200 %)[88]. This is a major reason ^{13}C spectra are acquired in the proton-decoupled mode. Another implication of nOe, the one that is useful for our conformational study, is that because of the nOe effect, NMR experiments can be designed and used to establish the spatial relationship between nuclei (e.g., protons) based on the change in absorption intensity. The nOe cross peaks in 2D NOESY spectra indicate the spatial separation between protons (nearer than approximately 4-5 Å give rise to nOe enhanced peaks) and

is due to spin polarization rather than spin coupling. Therefore, NOESY protons are near in space, but may also be separated by many intervening bonds. The intensity of the cross peak is dependent on the distance (r) between the interacting protons. In fact, the intensity enhancement is proportional to the inverse sixth power of the distance between the two nuclei ($1/r^{-6}$)[89]. By using a combination of NMR experiments (1D, COSY, HETCOR, NOESY, etc) the solution structure and conformation of macromolecules, such as nucleic acids, can be determined.

2.5.2. NMR Studies on B-Z Transition of DNA

In 1979, Patel, D.J. et al.[91] observed the conformational interconversion between two forms for poly(dG-dC) by using NMR. The 1D ^{1}H NMR was used to assign the cluster of chemical shifts of H1′, H3′, H5′, H8 of dG, and H5, H6 of dC. The data showed that the chemical shift of H1′ of dG increased at high salt concentration, suggesting that the glycosidic torsion angle of the purine bases have changed dramatically. The observation made by NMR was correlated to the conformational transition that had been previously detected by CD. Considering that the chemical shift of H1′ of dG is sensitive to glycosidic torsion angle, the NMR data were consistent with a conformation change of B-DNA to "alternating B-DNA", later has been identified as Z-DNA. Although the conformational change could be surmised from the simple 1D ^{1}H NMR of the oligonucleotides, the structural details could not be determined from the 1D spectrum which made full determination of the DNA conformational interconversion difficult to investigate. Due to the fact that nucleic acids are composed of nucleotide monomers, they contain many protons which have relatively similar chemical environments and hence

similar chemical shifts. Therefore, 1D proton NMR of DNA contains a series of absorption peaks that are not well resolved from one another and it is nearly impossible to assign or distinguish each proton based solely on chemical shift. The 2D correlation experiments, including COSY and NOESY, provided the possibility to assign all protons of nucleic acids and resulted in a more complete picture of structure and conformation. Hence, the 2D NMR experiments can be used to study the tertiary structure of nucleic acids, B and Z-DNA in particular.

For example, the sequential resonance assignment[92] based on NOESY and COSY experiment has been applied to the non-exchangeable proton assignment of the right handed helix of d(TGAGCGG) : d(CCGCTCA). The *J*-correlated spectroscopy based COSY, was used first to assign the thymine H6 and 5-methyl, cytosine H6 and H5, deoxyribose H1′, H2′, and H2″ resonances (see Figure 2.6 for structure of bases and nuclei number assignment). The off diagonal cross peaks in 2D nOe spectra link two types of protons that are near in space (less than 4 A°) and have been used to help identify the remaining adenine H8, guanine H8, H3′, H4′, H5′ and H5″ protons.

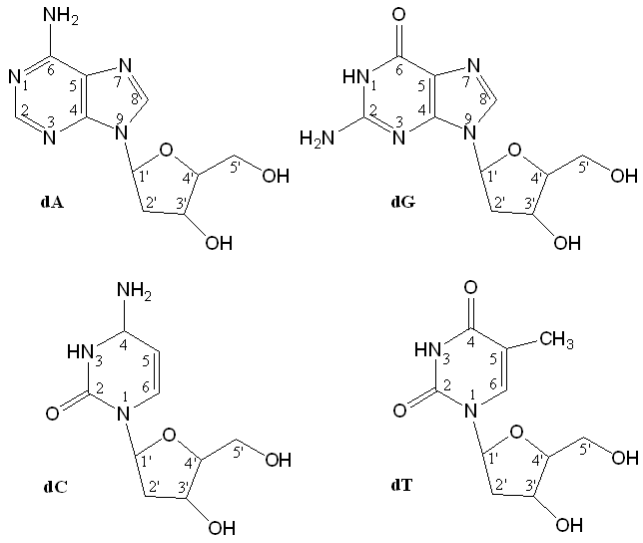


Figure 2.6 Structures of purine and pyrimidine deoxyribose nucleosides.

The nucleotides at 5′ or 3′-ends can serve as a starting point for the assignment due to distinctive chemical shift of the terminal bases compared to the same type of base in the middle of DNA strand (due to fraying of the ends). Then, the nearest neighbor nucleotides[93] can be targeted and assigned accordingly. Basically, one can start with a terminal base assignment and then use nOe and/or *J*-correlated cross peaks to assign the protons, sequentially, of the remaining bases until the entire sequence has been assigned.

The structural differences between Z-DNA and B-DNA stem from their distinct sugar backbone conformation. In B-DNA, all deoxyriboses are in *anti* conformation while Z-DNA has an alternating *syn-anti* backbone. These differences give rise to unique NMR spectra, mainly,due to the effect of conformation on the chemical shifts of the H1' protons and also the conformational requirements[94-96]. In the case of poly d(CG), the most notable observation is that the Z conformation displays strong nOe cross peak correlations between the guanine H8 and H1′[95]. This is due to the *syn* conformation of the purine nucleotides that brings the H8 of guanine into close proximity of H1′ of the intranucleoside deoxyribose residue (Figure 2.7). On the other hand, B-DNAs show only weak or no nOe cross peaks between the guanine H8 and its own H1′. This effect extends to C8 modified purines. For example, C8 methyl guanine containing DNA[20,21,97] displays strong nOe between methyl protons of the C8-methyl group and H1′.

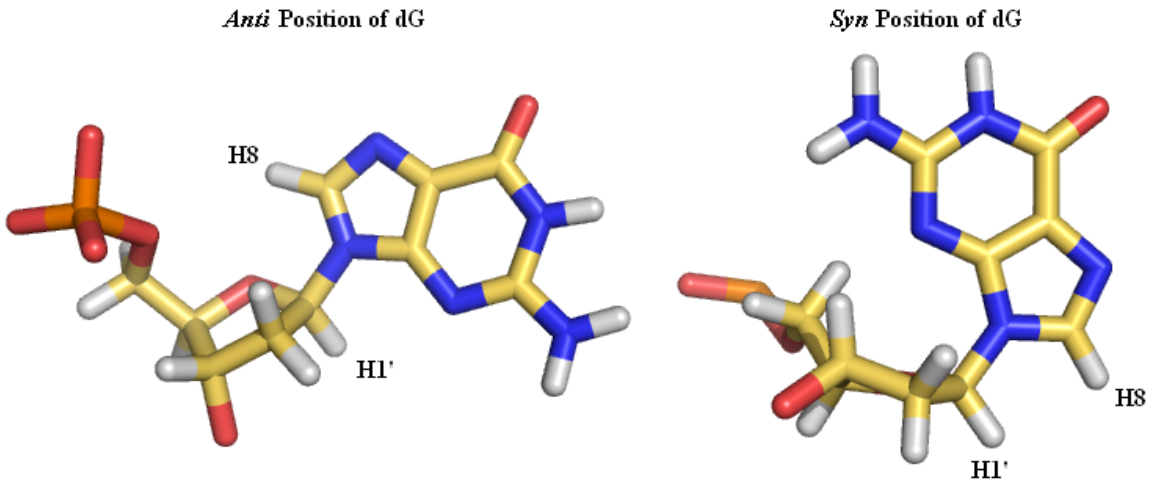


Figure 2.7 Shown are the *syn* and *anti* conformations of dG and how the H8 and H1′ are in close proximity in the *syn* conformation but not in the *anti* conformation.

By using a combination of NMR experiments as described above, the chemical shifts of protons of the unmodified CG and CG^{8Ph} have been assigned for both the B and Z conformation[22]. The structural information obtained from NMR is in complete agreement with the CD data and has confirmed that the C8-phenyl adduct facilitate the B-Z transition of CG decamers. The same NMR techniques have also been used in this research to study the B-Z transition of the rest of the aryl modified CG decamers.

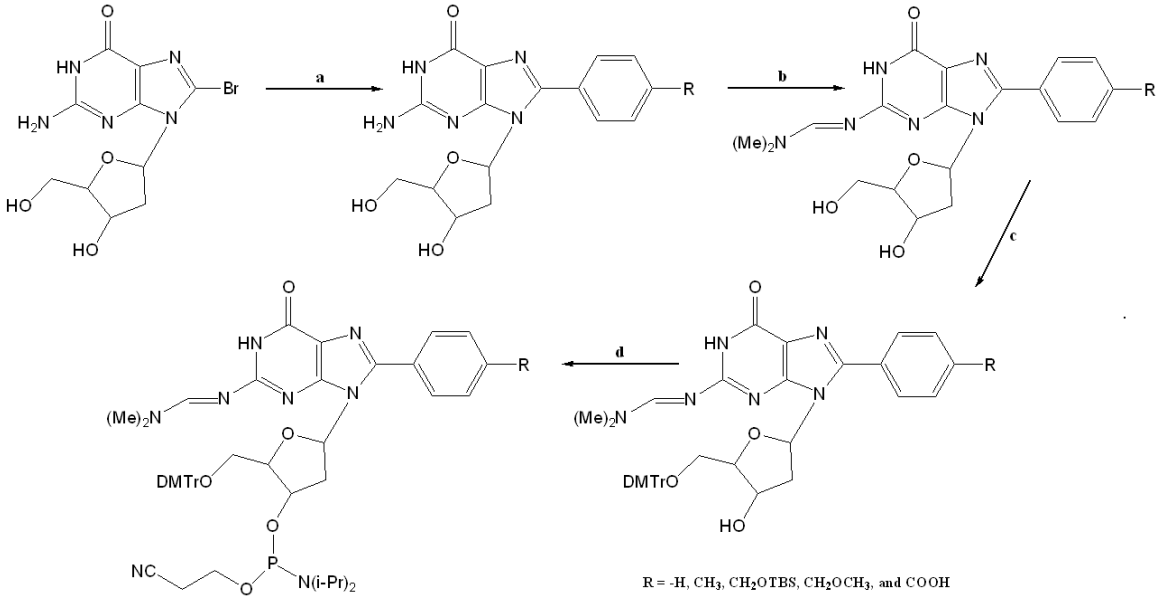
CHAPTER 3
EXPERIMENTAL PROCEDURES AND RESULTS PART I: SYNTHESIS OF CG DECAMERS CONTAINING C8-ARYL MODIFIED GUANINE

Our approach to the study of the effects of the C8-arylguanine adducts on B/Z-DNA equilibrium, was to prepare aryl modified CG decamers. Z-DNA formation generally requires an alternating purine-pyrimidine sequence. Thus, the modification could be incorporated into either an alternating TG or CG sequence. As it is well known, CG sequences are more prone to adopt the Z conformation[98], we used this sequence here. The aryl adducts selected for incorporation into this sequence were *p*-substituted (-H, -CH_3, -COO^-, -CH_2OCH_3, and -CH_2OH) and were selected because of the availability of carcinogenic profiles of the arylhydrazines that lead to their formation. Each adduct was introduced in the middle of the CG decamer sequence at C8 position of G6 base and since the CG sequences are palindromes, duplexes contain two modified guanines.

In our previous work[22], the C8 phenyl or C8-tolyl adducts, which are needed for phosphoramidite synthesis and hence oligonucleotide synthesis, were prepared via a Suzuki cross coupling between 8-BrdG and the corresponding aryl boronic acid. The C8-aryl modified dG phosphoramidites were then prepared by the successive protection of the N2 amine, the 5′-OH, and the 3′-OH groups. Finally, the unmodified CG, modified oligonucleotides CG^{8Ph}, and CG^{8Tol} were prepared by automated DNA synthesis, although the aryl modified bases were coupled by manual addition.

In the present case, the same basic methods were applied to prepare the modified oligonucleotides CG^{8CPh}, CG^{8MMPh}, CG^{8HMPh} as well as the synthesis of the corresponding modified phosphoramidites (Scheme 3.1) with one notable exception. The *p*-

hydroxymethyl substituent was found to interfere with DNA synthesis. Therefore, the *p*-hydroxymethyl was protected with TBS protecting group prior to Suzuki coupling to avoid potential side reactions during oligonucleotide synthesis. The TBS group was used to protect the hydroxymethyl group because it is stable toward all conditions used to prepare the phosphoramidite as well as conditions used for automated DNA synthesis and is removed under the same conditions used to remove all other protecting groups (i.e., concentrated ammonium hydroxide). Interestingly, although the *p*-carboxylate can be considered to be nucleophilic, the substituent did not interfere with DNA synthesis, hence no protection of the carboxyl group was necessary.



Scheme 3.1 Synthesis procedures of C8-arylguanine phosphoramidites a) Suzuki coupling, b) N2 protection, c) 5′-OH protection, and d) 3′-OH phosphoramidation

3.1 Synthesis of C8-Arylguanine Nucleosides and Phosphoramidites

3.1.1 General

All chemicals used were purchased from Aldrich (Milwaukee, WI) unless otherwise noted. Both 4-carboxyphenylboronic acid (4-CPBA) and 4-methoxymethylphenylboronic acid (4-MMPBA) were purchased from Frontier Scientific (Logan, Utah).

With exception for the Suzuki coupling reaction, all other reactions require anhydrous conditions. Therefore, organic solvents were dried and distilled prior to use as follows. Methanol was dried over molecular sieves (3Å). Methylene chloride was dried by distillation from phosphorus pentoxide. TEA, THF, and pyridine were dried and then distilled from LAH. Benzene and toluene were dried by distillation from calcium hydride. DMF was dried by distillation from barium oxide.

To identify and characterize the chemical structure of the synthesized nucleosides (Structures are shown in Appendix B), NMR spectra were obtained on either Varian 300 or 600 spectrometers (Palo Alto, CA). ^{1}H NMR and COSY were obtained for proton assignment, while ^{13}C NMR, HETCOR, and long range HETCOR were obtained to assign carbon nuclei. ESI-MS of the nucleosides, measured in positive ion mode, were recorded on a Finnigan LCQdeca (Waltham, MA). UV spectra were obtained on a Beckman DU640 spectrophotometer (Somerset, NJ). NMR, UV, and ESI-MS spectra of the modified nucleosides can be found in Appendix C.

3.1.2 Synthesis

8-Bromo-2′-deoxyguanosine, $C_{10}H_{12}O_4N_5Br$ (1): dG (2.86 g, 10.0 mmol) was suspended in the mixture of acetonitrile (100 ml) and water (25 ml) and N-bromosuccinimide (freshly recrystallized, NBS, 2.67 g, 15.0 mmol) was then added to the suspension and the formation of light orange precipitate was then observed. The reaction mixture was allowed to stir for 30 min at room temperature. The precipitate was isolated by filtration and then suspended in acetone (50 ml). The suspension was stirred for an additional 2 h at room temperature, cooled at -20°C overnight, the precipitate collected by filtration, and then dried *in vacuo* to yield 8-BrdG (2.81 g, 8.1 mmole, 81.2% yield).

^{1}H NMR (600 MHz, dmso-d6): δ ppm 10.78 (1H, s, NH), 6.47 (2H, s, NH_2), 6.16 (1H, t, *J* = 7.5 Hz, H-1′), 5.23 (1H, d, *J* = 4.2 Hz, 3′-OH), 4.83 (1H, t, J = 5.7 Hz, 5′-OH), 4.39 (1H, m, H-3′), 3.80 (1H, m, H-4′), 3.62 and 3.50 (2H, m, H-5′/5″), 3.16 (1H, m, H-2″) and 2.10 (1H, m, H-2′). ^{13}C NMR (150 MHz, dmso-d6): δ ppm 155.42, 153.31, 151.97, 120.50, 117.49, 87.89, 85.06, 71.02, 62.04, and 36.47.

4-(TBS-O-methyl)-phenyl bromide, $C_{13}H_{21}OSiBr$ (2): 4-Bromobenzyl alcohol (756 mg, 4.0 mmol), TBSCl (1.30 mg, 8.0 mmol), and imidazole (605 mg, 8.8 mmol) were dissolved in DMF (10 mL). The reaction mixture was allowed to stir at room temperature for 24 hr under argon. The reaction mixture was then concentrated to dryness using rotavap (vacuum pump). The crude product was extracted with hexane (30 mL), washed

with water (3 x 30 mL), and the organic layer dried over sodium sulfate to yield **2** (1.08 g, 3.58 mmol, 89.6%) was then collected by filtration and brought to dryness *in vacuo.*

^{1}H NMR (600 MHz, dmso-d6): δ ppm 7.46 (2H, d, *J* = 8.4 Hz, aryl), 7.20 (2H, d, *J* = 9.0 Hz, aryl), 4.69 (2H, s, CH_2), 0.95 (9H, s, t-butyl), and 0.11 (6H, s, dimethylsilyl). ^{13}C NMR (150 MHz, dmso-d6): δ ppm 140.70, 131.49, 127.94, 120.80, 64.55, 26.14, 18.60, and -5.02.

4-(TBS-O-methyl)-phenyl boronic acid, $C_{13}H_{23}O_3BSi$ (3): The synthesis protocol has been adapted based on the previous report by Zheng et al.[70] Compound **(2)** (1 g, 3.3 mmol) was dissolved in THF (10 mL), toluene (2 mL), and triisopropyl borate (1.1 mL, 4.7 mmol). The mixture was sparged with nitrogen gas at -78°C (acetone/dry ice bath) for 30 min and then *n*-butyl lithium (1.5 M in hexane, 4mL) was gradually added to the solution over 2 hr The reaction mixture was quenched with 2M HCl (3.5 mL) in an ice bath and stirred for 30 min. Ethyl acetate (5.5 mL) was then added to the reaction flask and mixed for an additional 30 min. The organic layer was collected, washed with the mixture of 5% sodium bicarbonate solution (5 mL) and sodium chloride (450 mg), and concentrated to dryness *in vacuo*. The crude product was suspended in ACN and the precipitate was collected by filtration to yield the aryl boronic acid **(3)** (300 mg, 1.13 mmol, 33.7 %).

^{1}H NMR (600 MHz, dmso-d6): δ ppm 8.22 (2H, d, *J* = 8.4 Hz, aryl), 7.48 (2H, d, *J* = 7.8 Hz, aryl), 4.86 (2H, s, CH_2), 0.99 (9H, s, t-butyl), and 0.15 (6H, s, dimethylsilyl). ^{13}C

NMR (150 MHz, dmso-d6): δ ppm 146.52, 135.90, 133.67, 125.66, 65.19, 26.18, 18.66, and -4.98.

8-(4-Carboxyphenyl)-2′-deoxyguanosine, $C_{17}H_{17}O_6N_5$ (4): 4-CPBA (183 mg, 1.10 mmol), 8-BrdG (346 mg, 1.00 mmol), TPPTS (80 mg, 0.14 mmol), and Na_2CO_3 (292 mg, 2.75 mmol) were placed in round bottom flask and purged with argon for 10 min. Pd(II)acetate (22 mg, 0.10 mmol) was then added to the flask and again purged with argon for an additional 5 min. A mixture of acetonitrile and water (1:2, 10 mL) (sparged with argon) was then added to the reaction flask and the reaction mixture then heated at 80ºC, under argon, and allowed to react until RP-TLC (C18, 1:1 water:methanol, UV detection) indicated complete consumption of 8-BrdG. The reaction mixture was then diluted with water (5 mL) and the pH was adjusted to 6 by the dropwise addition of 10 % HCl. The precipitate was collected and dried *in vacuo* to yield the coupled product (320 mg, 0.83 mmol, 82.6% yield).

^{1}H NMR (600 MHz, dmso-d6): δ ppm 10.82 (1H, s, NH), 8.07 (2H, d, J = 7.8 Hz, aryl), 7.80 (2H, d, J = 8.4 Hz, aryl), 6.48 (2H, s, NH_2), 6.10 (1H, t, J = 7.2 Hz, H-1′), 5.14 (1H, brs, 3′-OH), 4.96 (1H, brs, 5′-OH), 4.34 (1H, m, H-3′), 3.80 (1H, m, H-4′), 3.66 and 3.55 (2H, m, H-5′/5″), 3.11 (1H, m, H-2″) and 2.05 (1H, ddd, J = 3.0, 6.6, 13.2 Hz, H-2′).^{13}C NMR (150 MHz, dmso-d6): δ ppm 166.88, 156.64, 153.20, 152.29, 146.11, 134.33, 131.36, 129.48, 129.24, 117.42, 87.85, 84.45, 71.00, 61.93, and 36.66. UV: ε^{227} = 13,745 $cm^{-1}M^{-1}$ and ε^{285} = 18,054 $cm^{-1}M^{-1}$. MS: m/z for MW 387.35, calculated MH^+ 388.35, found 388, 410 $(M+Na)^+$, 426 $(M+K)^+$, 775 $(2M+H)^+$,and 797 $(2M+Na)^+$.

8-(4-Methoxymethylphenyl)-2′-deoxyguanosine, $C_{18}H_{21}O_5N_5$ (5): 4-MMPBA (183 mg, 1.10 mmol), 8-BrdG (346 mg, 1.00 mmol), TPPTS (80 mg, 0.14 mmol), and Na_2CO_3 (292 mg, 2.75 mmol) were placed in round bottom flask and purged with argon for 10 min. Pd(II)acetate (22 mg, 0.10 mmol) was then added to the flask and again purged with argon for an additional 5 min. A mixture of acetonitrile and water (1:2, 10 mL) (sparged with argon) was then added to the reaction flask and the reaction mixture then heated at 80ºC, under argon, and allowed to react until RP-TLC (C18, 1:1 water:methanol, UV detection) indicated complete consumption of 8-BrdG. The reaction mixture was then diluted with water (5 mL) and the pH was adjusted to 6 by the dropwise addition of 10 % HCl. The precipitate was collected and dried *in vacuo* to yield the coupled product (260 mg, 0.67 mmol, 67.1% yield).

^{1}H NMR (600 MHz, dmso-d6): δ ppm 10.76 (1H, s, NH), 7.63 (2H, d, J = 8.4 Hz, aryl), 7.47 (2H, d, J = 7.8 Hz, aryl), 6.39 (2H, s, NH_2), 6.07 (1H, dt, J = 1.8, 7.5 Hz, H-1′), 5.12 (1H, d, J = 4.2 Hz, 3′-OH), 4.97 (1H, t, J = 5.7 Hz, 5′-OH), 4.49 (2H, s, CH_2), 4.33 (1H, m, H-3′), 3.79 (1H, m, H-4′), 3.65 and 3.54 (2H, m, H-5′/5″), 3.34 (3H, s, OCH_3), 3.15 (1H, m, H-2″) and 2.02 (1H, m, H-2′). ^{13}C NMR (150 MHz, dmso-d6): δ ppm 156.63, 152.97, 151.95, 146.92, 139.69, 129.37, 129.05, 127.48, 117.13, 87.84, 84.58, 73.14, 71.15, 62.07, 57.71, and 36.55. UV: ε^{221} = 9,377 $cm^{-1}M^{-1}$ and ε^{282} = 15,552 $cm^{-1}M^{-1}$. MS: m/z for MW 387.40, calculated MH^+ 388.40, found 411 $(M+Na)^+$ and 798 $(2M+Na)^+$.

8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine, $C_{23}H_{33}O_5N_5Si$ (6): Compound **(3)** (293 mg, 1.10 mmol), 8-BrdG (346 mg, 1.00 mmol), TPPTS (80 mg, 0.14 mmol), and

Na_2CO_3 (292 mg, 2.75 mmol) were placed in round bottom flask and purged, with argon, for 10 min. Pd(II)acetate (22 mg, 0.10 mmol) was then added to the flask and again purged with argon for an additional 5 min. A mixture of acetonitrile and water (1:1, 10 mL) (sparged with argon) was then added to the reaction flask and the reaction mixture then heated at 80ºC, under argon, and allowed to react until RP-TLC (C18, 1:1 water:methanol, UV detection) indicated complete consumption of 8-BrdG. The reaction mixture was then diluted with water (5 mL). Unlike the previous coupling reaction, the pH of reaction mixture was not adjusted. The precipitate was collected, washed with water and ethyl acetate respectively, and then dried *in vacuo* to yield the coupled product (350 mg, 0.72 mmol, 71.8% yield).

^{1}H NMR (600 MHz, dmso-d6): δ ppm 10.78 (1H, s, NH), 7.62 (2H, d, J = 8.4 Hz, aryl), 7.46 (2H, d, J = 8.4 Hz, aryl), 6.39 (2H, s, NH_2), 6.06 (1H, t, J = 7.5 Hz, H-1′), 5.12 (1H, d, J = 3.6 Hz, 3′-OH), 4.96 (1H, m, 5′-OH), 4.80 (2H, s, CH_2), 4.34 (1H, m, H-3′), 3.79 (1H, m, H-4′), 3.65 and 3.55 (2H, m, H-5′/5″), 3.16 (1H, m, H-2″), 2.02 (1H, m, H-2′), 0.93 (9H, s, t-butyl), and 0.11 (6H, s, dimethylsilyl). ^{13}C NMR (150 MHz, dmso-d6): δ ppm 156.71, 153.00, 151.91, 147.00, 142.54, 129.00, 128.86, 125.99, 117.13, 87.87, 84.65, 71.20, 63.90, 62.12, 36.56, 25.83, 18.01, and -5.33. UV: ε^{281} = 15,285 $cm^{-1}M^{-1}$. MS: m/z for MW 487.63, calculated MH^+ 488.63, found 488, 510 $(M+Na)^+$, 526 $(M+K)^+$, 975 $(2M+H)^+$, and 997 $(2M+Na)^+$.

N2-(N,N-Dimethylformamidine)-8-(4-carboxyphenyl)-2′-deoxyguanosine, $C_{20}H_{22}O_6N_6$ (7): Compound **(4)** (387 mg, 1.00 mmol) was dissolved in methanol (10

mL). N,N-dimethylformamide dimethyl acetal (0.67 mL, 5.00 mmol) was then added and the reaction mixture was allowed to stand for 24 h, under argon, at room temperature. The reaction mixture was then concentrated to dryness *in vacuo* to yield product **(7)** (423 mg, 0.95 mmole, 95.7 % yield).

^{1}H NMR (600 MHz, dmso-d6): δ ppm 8.53 (1H, s, HC=N), 8.02 (2H, d, *J* = 8.4 Hz, aryl), 7.64 (2H, d, *J* = 8.4 Hz, aryl), 6.14 (1H, t, *J* = 7.2 Hz, H-1′), 4.45 (1H, m, H-3′), 3.82 (1H, m, H-4′), 3.68 and 3.58 (2H, m, H-5′/5″), 3.18 (1H, m, H-2″), 3.16 and 3.05 (3H each, s, $N(CH_3)_2$), and 2.10 (1H, ddd, *J* = 3.0, 6.6, 13.2 Hz, H-2′).^{13}C NMR (150 MHz, dmso-d6): δ ppm 169.17, 158.12, 156.82, 150.75, 147.98, 138.74, 131.22, 129.13, 128.46, 120.21, 87.67, 84.71, 70.95, 61.94, 40.79, 37.08, 34.58, and 34.13. UV: ε^{231} = 17,153 $cm^{-1}M^{-1}$ and ε^{318} = 20,164 $cm^{-1}M^{-1}$. MS: m/z for MW 442.43, calculated MH^+ 443.43, found 443, 465 $(M+Na)^+$, 488 $(M+2Na)^+$, 885 $(2M+H)^+$, and 907 $(2M+Na)^+$.

N2-(N,N-Dimethylformamidine)-8-(4-methoxymethylphenyl)-2′-deoxyguanosine, $C_{21}H_{26}O_5N_6$ (8): Compound **(5)** (387 mg, 1.00 mmol) was dissolved in methanol (10 mL). N,N-dimethylformamide dimethyl acetal (0.67 mL, 5.00 mmol) was then added and the reaction mixture was allowed to stand for 24 h, under argon, at room temperature. The reaction mixture was concentrated down close to dryness using a Rotavap follow by the addition of water (2 ml). After filtration, the precipitate was dried *in vacuo* to yield product **(8)** (388 mg, 0.88 mmole, 87.6 % yield).

^{1}H NMR (600 MHz, dmso-d6): δ ppm 11.45 (1H, s, NH), 8.51 (1H, s, HC=N), 7.64 (2H, d, *J* = 8.4 Hz, aryl), 7.49 (2H, d, *J* = 7.8 Hz, aryl), 6.12 (1H, t, *J* = 7.2 Hz, H-1′), 5.21 (1H, d, *J* = 4.2 Hz, 3′-OH), 4.88 (1H, dt, *J* = 1.8, 6.0 Hz, 5′-OH), 4.50 (2H, s, CH_2), 4.44 (1H, m, H-3′), 3.82 (1H, m, H-4′), 3.68 and 3.57 (2H, m, H-5′/5″), 3.34 (3H, s, OCH_3), 3.22 (1H, m, H-2″), 3.16 and 3.05 (3H each, s, $N(CH_3)_2$), and 2.08 (1H, m, H-2′).^{13}C NMR (150 MHz, dmso-d6): δ ppm 158.10, 157.48, 156.81, 150.67, 147.92, 139.86, 129.19, 129.08, 127.56, 120.17, 87.70, 84.81, 73.13, 71.03, 61.99, 57.72, 40.78, 37.03, and 34.57. UV: ε^{229} = 14,602 $cm^{-1}M^{-1}$ and ε^{314} = 20,834 $cm^{-1}M^{-1}$. MS: m/z for MW 442.47, calculated MH^+ 443.47, found 443, 466 $(M+Na)^+$, and 908 $(2M+Na)^+$.

N2-(N,N-Dimethylformamidine)-8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine, $C_{26}H_{38}O_5N_6Si$ (9): Compound **(6)** (488 mg, 1.00 mmol) was dissolved in methanol (10 mL). N,N-dimethylformamide dimethyl acetal (0.67 mL, 5.00 mmol) was then added and the reaction mixture was allowed to stand for 24 h, under argon, at room temperature. The reaction mixture was concentrated to near dryness using Rotavap follow by the addition of water (2 ml). After filtration, the precipitate was dried *in vacuo* to yield product **(9)** (419 mg, 0.77 mmole, 77.3 % yield).

^{1}H NMR (600 MHz, dmso-d6): δ ppm 11.42 (1H, s, NH), 8.51 (1H, s, HC=N), 7.63 (2H, d, *J* = 7.8 Hz, aryl), 7.48 (2H, d, *J* = 7.8 Hz, aryl), 6.11 (1H, t, *J* = 7.2 Hz, H-1′), 5.20 (1H, d, *J* = 2.4 Hz, 3′-OH), 4.88 (1H, m, 5′-OH), 4.81 (2H, s, CH_2), 4.44 (1H, m, H-3′), 3.82 (1H, m, H-4′), 3.67 and 3.57 (2H, m, H-5′/5″), 3.22 (1H, m, H-2″), 3.16 and 3.05 (3H each, s, $N(CH_3)_2$), 2.08 (1H, m, H-2′), 0.93 (9H, s, t-butyl), and 0.12 (6H, s,

dimethylsilyl). ^{13}C NMR (150 MHz, dmso-d6): δ ppm 158.07, 157.46, 156.77, 150.62, 148.01, 142.70, 129.02, 128.67, 126.03, 120.15, 87.70, 84.83, 71.04, 63.88, 62.01, 40.76, 37.03, 34.56, 25.82, 18.00, and -5.35. UV: ε^{229} = 21,603 $cm^{-1}M^{-1}$ and ε^{313} = 28,129 $cm^{-1}M^{-1}$. MS: m/z for MW 542.71, calculated MH^+ 543.71, found 543, 565 $(M+Na)^+$, 1085 $(2M)^+$, and 1107 $(2M+Na)^+$.

5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-carboxyphenyl)-2′-deoxyguanosine, $C_{41}H_{40}O_8N_6$ (10): Compound **(7)** (442 mg, 1.00 mmol) and DMTr-Cl (1 g, 2.80 mmol) were dissolved in pyridine (10 mL) and then TEA (430 µL, 3.08 mmol) was added. The reaction mixture was stirred at room temperature, under argon, for 1 h. The reaction was quenched by the addition of methanol (10 mL) and solvents then removed *in vacuo*. The crude product was resuspended in methylene chloride (100 mL) and stored at -20°C for 24 hr. The precipitate was collected and dried again *in vacuo* to yield the DMTr protected product **(10)** (230 mg, 0.31 mmole, 30.9 % yield).

^{1}H NMR (600 MHz, dmso-d6): δ ppm 8.32 (1H, s, HC=N), 8.05 (2H, d, *J* = 7.8, aryl), 7.83 (2H, d, *J* = 8.4, aryl), 7.32-6.73 (13H, m, DMTr-H), 6.20 (1H, dd, *J* = 4.8, 7.8, H-1′), 4.57 (1H, m, H-3′), 3.93 (1H, m, H-4′), 3.71 and 3.70 (3H each, s, OCH_3), 3.30 and 3.15 (2H, m, H-5′/5″), 3.21 (1H, m, H-2″), 3.03 and 2.98 (3H each, s, $N(CH_3)_2$), and 2.20 (1H, m, H-2′). ^{13}C NMR (150 MHz, dmso-d6): δ ppm 157.93, 157.86, 157.59, 157.50, 156.44, 150.54, 147.53, 144.92, 135.67, 135.58, 129.59, 129.41, 129.31, 128.95, 127.62, 127.56, 126.48, 120.28, 112.93, 112.90, 85.71, 85.18, 84.05, 70.82, 63.99, 54.93, 54.89, 40.74,

37.44, 34.60, and 34.18. UV: ε^{233} = 43,258 $cm^{-1}M^{-1}$ and ε^{322} = 27,453 $cm^{-1}M^{-1}$. MS: m/z for MW 744.80, calculated MH^+ 745.80, found 745, 767 $(M+Na)^+$, and 1511 $(2M+Na)^+$.

5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-methoxymethylphenyl)-2′-deoxyguanosine $C_{42}H_{44}O_7N_6$ (11): Compound **(8)** (442 mg, 1.00 mmol) and DMTr-Cl (535 mg, 1.50 mmol) were dissolved in pyridine (10 mL) and then TEA (229 µL, 1.65 mmol) was added. The reaction mixture was stirred at room temperature, under argon, and was monitored by normal phase TLC (aluminum oxide, 6 % methanol in methylene chloride) until the complete consumption of **(8)** was observed. In addition, an ethanolic solution of phosphomolybdic acid (10%) was used on the developed TLC plate to visualize the production of **(11)**. Upon completion of the reaction or after 4 hr, the reaction was quenched by the addition of methanol (10 mL) and solvents removed *in vacuo*. The crude product was then purified by low pressure column chromatography (aluminum oxide, 0-10% methanol in methylene chloride) to yield the DMTr protected product **(11)** (406 mg, 0.55 mmole, 54.5 % yield).

^{1}H NMR (600 MHz, dmso-d6): δ ppm 11.42 (1H, s, NH), 8.31 (1H, s, HC=N), 7.74 (2H, d, *J* = 8.4, aryl), 7.46 (2H, d, *J* = 8.4, aryl), 7.33-6.73 (13H, m, DMTr-H), 6.17 (1H, dt, *J* = 3.0, 6.6, H-1′), 5.28 (1H, d, *J* = 4.8 Hz, 3′-OH), 4.57 (1H, m, H-3′), 4.50 (2H, s, CH_2), 3.92 (1H, m, H-4′), 3.71 and 3.70 (3H each, s, OCH_3), 3.35 (3H, s, CH_3), 3.30 and 3.15 (2H, m, H-5′/5″), 3.20 (1H, m, H-2″), 3.03 and 2.98 (3H each, s, $N(CH_3)_2$), and 2.18 (1H, m, H-2′). ^{13}C NMR (150 MHz, dmso-d6): δ ppm 157.94, 157.87, 157.52, 156.31, 150.41, 148.05, 144.94, 139.75, 135.70, 135.59, 129.61, 129.42, 129.12, 127.62, 127.57, 127.50,

126.48, 120.07, 112.94, 112.91, 85.72, 85.18, 84.02, 73.14, 70.87, 64.04, 57.68, 54.94, 54.90, 40.72, 37.37, and 34.59. UV: ε^{231} = 35,034 $cm^{-1}M^{-1}$ and ε^{319} = 23,448 $cm^{-1}M^{-1}$. MS: m/z for MW 744.85, calculated MH^+ 745.85, found 746, 768 $(M+Na)^+$, 784 $(M+K)^+$, and 1512 $(2M+Na)^+$.

5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine, $C_{47}H_{56}O_7N_6Si$ (12): Compound **(9)** (442 mg, 1.00 mmol) and DMTr-Cl (535 mg, 1.50 mmol) were dissolved in pyridine (10 mL) and then TEA (229 µL, 1.65 mmol) was added. The reaction mixture was stirred at room temperature, under argon, and was monitored by normal phase TLC (aluminum oxide, 6 % methanol in methylene chloride) until the complete consumption of **(9)** was observed. An ethanolic solution of phosphomolybdic acid (10%) was used on the developed TLC plate to visualize the production of **(12)**. Upon completion of the reaction or after 4 hr, the reaction was quenched by the addition of methanol (10 mL) and solvents then removed *in vacuo*. The crude product was then purified by low pressure column chromatography (aluminum oxide, 0-10% methanol in methylene chloride) to yield the DMTr protected product **(12)** (350 mg, 0.41 mmole, 41.4 % yield).

^{1}H NMR (600 MHz, dmso-d6): δ ppm 11.41 (1H, s, NH), 8.31 (1H, s, HC=N), 7.74 (2H, d, *J* = 8.4, aryl), 7.46 (2H, d, *J* = 7.8, aryl), 7.33-6.73 (13H, m, DMTr-H), 6.17 (1H, m, H-1′), 5.28 (1H, d, *J* = 4.8 Hz, 3′-OH), 4.81 (2H, s, CH_2), 4.57 (1H, m, H-3′), 3.93 (1H, m, H-4′), 3.71 and 3.70 (3H each, s, OCH_3), 3.32 and 3.15 (2H, m, H-5′/5″), 3.21 (1H, m, H-2″), 3.03 and 2.98 (3H each, s, $N(CH_3)_2$), 2.18 (1H, m, H-2′), 0.94 (9H, s, t-butyl), and

0.12 (6H, s, dimethylsilyl). ^{13}C NMR (150 MHz, dmso-d6): δ ppm 157.93, 157.86, 157.51, 156.27, 150.38, 148.16, 144.94, 142.61, 135.68, 135.59, 129.61, 129.41, 129.07, 128.93, 127.62, 127.56, 126.47, 125.97, 120.04, 113.09, 112.93, 112.89, 85.70, 85.17, 84.00, 70.86, 64.03, 63.89, 54.92, 54.89, 40.72, 37.37, 34.58, 25.81, 17.99, and -5.36. UV: ε^{231} = 39,719 $cm^{-1}M^{-1}$ and ε^{316} = 22,703 $cm^{-1}M^{-1}$. MS: m/z for MW 845.08, calculated MH^+ 846.08, found 845, 867 $(M+Na)^+$, and 1712 $(2M+Na)^+$.

3′-O-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-5′-O-(DMTr)-N2-(N,N dimethylformamidine)-8-(4-carboxyphenyl)-2′-deoxyguanosine, $C_{50}H_{57}O_9N_8P$ (13): Compound **(10)** (0.25 mmol) was dissolved in methylene chloride (2.5 mL) and then TEA (115 μL, 0.83 mmol) was added followed by 2-cyanoethyl diisopropylchlorophosphoamidite (100 μL, 0.43 mmol). After stirring the reaction mixture for 30 min at room temperature, another portion of 2-cyanoethyl diisopropylchlorophosphoamidite (75 μL, 0.32 mmol) was added to the reaction mixture and then stirred for an additional 30 min. The crude reaction was dried *in vacuo.* The reaction mixture was resuspended in a mixture of benzene and THF (4:1, 3 mL), filtered and the filtrate was concentrated to dryness *in vacuo* to yield the crude phosphoramidite **(13)** that is ready to use for DNA synthesis.

3′-O-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-5′-O-(DMTr)-N2-(N,N dimethylformamidine)-8-(4-methoxymethylaryl)-2′-deoxyguanosine, $C_{51}H_{61}O_8N_8P$ (14): Compound **(11)** (0.25 mmol) was dissolved in methylene chloride (2.5 mL) and then TEA (115 μL, 0.83 mmol) was added followed by 2-cyanoethyl

diisopropylchlorophosphoamidite (100 µL, 0.43 mmol). After stirring the reaction mixture for 30 min at room temperature, another portion of 2-cyanoethyl diisopropylchlorophosphoamidite (75 µL, 0.32 mmol) was added to the reaction mixture and then stirred for an additional 30 min. The crude reaction was dried *in vacuo.* The reaction mixture was resuspended in a mixture of benzene and THF (4:1, 3 mL), filtered and the filtrate was concentrated to dryness *in vacuo* to yield the crude phosphoramidite **(14)** for DNA synthesis.

3′-O-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-5′-O-(DMTr)-N2-(N,N dimethylformamidine)-8-(4-(TBS-O-methyl)aryl)-2′-deoxyguanosine, $C_{56}H_{73}O_8N_8SiP$ (15): Compound **(12)** (0.25 mmol) was dissolved in methylene chloride (2.5 mL) and then TEA (115 µL, 0.83 mmol) was added followed by 2-cyanoethyl diisopropylchlorophosphoamidite (100 µL, 0.43 mmol). After stirring the reaction mixture for 30 min at room temperature, another portion of 2-cyanoethyl diisopropylchlorophosphoamidite (75 µL, 0.32 mmol) was added to the reaction mixture and then stirred an for additional 30 min. The crude reaction was dried *in vacuo.* The reaction mixture was resuspended in a mixture of benzene and THF (4:1, 3 mL), filtered and the filtrate was concentrated to dryness *in vacuo* to yield the crude phosphoramidite **(15)** that is ready to use for DNA synthesis.

3.2 Synthesis of CG Decamers Containing C8-Aryl Modified Guanine

3.2.1 General

The oligonucleotides CG^{8CPh}, CG^{8MMPh}, and CG^{8HMPh} were made on ABI 394 DNA synthesizer (Newark, CA). The reagents used for automated DNA synthesis, including phosphoramidites (dG CE-PA and dC CE-PA), phosphoramidite diluent (anhydrous ACN), activator (0.45M tetrazole in ACN), cap mix A (THF/pyridine/acetic anhydride), cap mix B (16% 1-methylimidazole in THF), oxidizing solution (0.02M I_2 in THF/pyridine/water), deblocking mix (3% TCA in methelene chloride), and dG-CPG synthesis column (1 μmol, 500Å pore size), were purchased from Glen Research (Sterling, VA). In addition, the anhydrous acetonitrile used with the DNA synthesizer was purchased from VWR (West Chester, PA).

To purify the synthesized oligonucleotides, the following FPLC system has been used. Anion exchange columns (DEAE-5PW) were purchased from Tosoh Bioscience (Montgomeryville, PA). Waters U6K loop injector, HPLC pump model 510 or 501, and Waters 486 tunable absorbance detector (Milford, MA) were used for the chromatography. Sep Pak C18 cartridges for solid phase extraction used to desalt the DNA sample were bought from VWR (West Chester, PA). The amount of DNA in synthesized samples were quantified by UV measurement at 260 nm on a Beckman DU640 spectrophotometer (Somerset, NJ) and assumed that a 1 mL solution with an absorbance of 1 OD at 260 nm contained 24 μg of oligonucleotide duplex. ESI-MS spectra of the purified oligonucleotides, obtained in negative ion mode, were recorded on a Finnigan LCQdeca (Waltham, MA).

3.2.2 Synthesis

The modified oligonucleotides were made on an automated DNA synthesizer through solid phase synthesis on 1 μmol scale column. Fresh reagents were used to maximize the efficiency of the reactions. The base extensions started from the 3′ end through the 5′ end of 5′-CGCGCG*CGC-3′ sequences while the dG phosphoramidite from the dG-CPG synthesis column served as a first 3′-base. Generally, the synthesis utilized the standard protocol (1.0 μM CE cycle) for automated DNA synthesis which includes detritylation, coupling, capping and oxidation. The efficiency of each base extension can be monitored from the trityl group (as the alcohol) that is cleaved just before the addition of the next base. In order to make the aryl modified oligonucleotides, the G6 coupling was performed manually with the C8-aryl modified guanine phosphoramidites (compounds **(13)**, **(14)**, or **(15)** to yield CG^{8CPh}, CG^{8MMPh}, or CG^{8HMPh} respectively). Manual addition of the modified bases was required because these phosphoramidites form gels in acetonitrile within minutes of initial dissolution. After detritylation of C7 (the 3rd base from 3′end) the automated synthesis was stopped and the synthesis column, sealed at both ends to minimize air and moisture exposure, was removed from the DNA synthesizer. The C8-aryl modified dG phosphoramidite (25-30 mg or approximately 25 μmol) was dissolved immediately prior to use in anhydrous ACN (100 μL) and transferred to the synthesis column through a 1 mL syringe. The activator solution (100 μL) was then added to the synthesis column through another 1 mL syringe. Both phosphoramidite solution and activator solution were mixed through the synthesis column and the reaction mixture was allowed to stay in the column for 15 min. After that, the reaction mixture was removed from the column and a fresh solution of the

coupling reaction mixture was then added and allowed to stay in the column for an additional 15 min to maximize the efficiency of modified base extension. After removing the reaction mixture, the synthesis column was reattached to the DNA synthesizer. The automated synthesis cycle was then resumed after programming to skip the G6 coupling and pass to the capping step. The DNA synthesis was continued normally after the manual coupling of C8-aryl G6 until the CG decamer sequence has been completely extended.

Following the completion of DNA synthesis, the oligonucleotides were cleaved off the synthesis column. Ammonium hydroxide (30%, 0.3 mL) was transferred by using a syringe and allowed to stay on the column for 20 min. Two additional fresh fractions of ammonium hydroxide were added to the synthesis column in the same fashion as the first one to ensure complete removal of the oligonucleotides from the column. The synthesized oligonucleotides, in ammonium hydroxide, were then collected from the column, transferred to a centrifuge tube, and deprotected by heating at 55°C overnight. After cooling to room temperature, the solution was brought to dryness in SpeedVac to yield the crude C8-arylguanine modified CG decamer.

3.2.3 FPLC Purification

Generally, the crude oligonucleotide sample contains unwanted products, including incomplete extended sequences and residual protecting groups. Therefore, purification by FPLC is necessary to get a pure oligonucleotide for further experiments. The crude oligonucleotides (0.3-0.4 μmol) were dissolved in deionized (DI) water (250 μL) and were purified by FPLC using a gradient of 70-30% buffer A (10 mM NaOH) and

buffer B (30-70% 10 mM NaOH and 1 M NaCl) at a flow rate 6 mL/min on a Biorad TSK gel DEAE-5PW preparative column (21.5 mm ID x 15 cm). The separation of the oligonucleotides was monitored by UV detector at 260 nm and the mobile phase fractions from the major peaks on the chromatogram were collected due to the likelihood that the desirable DNA sequence is often the major product which gives the highest absorbance.

The fractions collected from FPLC have high salt concentration. Thus, a C18 Sep Pak was used to desalt the oligonucleotide sample. A solid phase extraction started by using a syringe to purge methanol (3 x 2ml) and water (3 x 2 ml) through the C18 cartridge respectively. The product fraction from FPLC was then slowly infused through the cartridge. At this point, the oligonucleotide would adsorb to the C18 matrices while the salt will pass through with solvent. Water (4 x 2mL) was later used to wash the remaining salts out of the cartridge. The oligonucleotide was collected from the cartridge after eluting with aqueous methanol (60% methanol, 4 x 1.5 mL). The solvent was removed in a SpeedVac to yield a pure oligonucleotide sample.

To confirm that the purity of oligonucleotide after chromatography and desalting, the sample has been analyzed using FPLC. The oligonucleotides sample (approximately 10 μM in DI water) was run with FPLC using a gradient of 70-30% buffer A (10 mM NaOH) and buffer B (30-70% 10 mM NaOH and 1 M NaCl) at a flow rate 1 mL/min on a Biorad TSK gel DEAE-5PW analytical column (5.0 mm ID x 5 cm). To eliminate the false result that a single peak is composed of more than one product, which may occur due to the binding of single-stranded oligonucleotides, the sample was heated at 90°C for 30 min to denature all DNA into single strand. The chromatogram obtained at 260 nm, that displayed a single peak, indicates that the sample contained only one oligonucleotide

which was successfully purified. The observed retention time of CG^{8CPh}, CG^{8MMPh}, and CG^{8HMPh} is 21.9, 17.7, and 25.3 min respectively as shown in the chromatograms in Figure 3.1, 3.2, and 3.3.

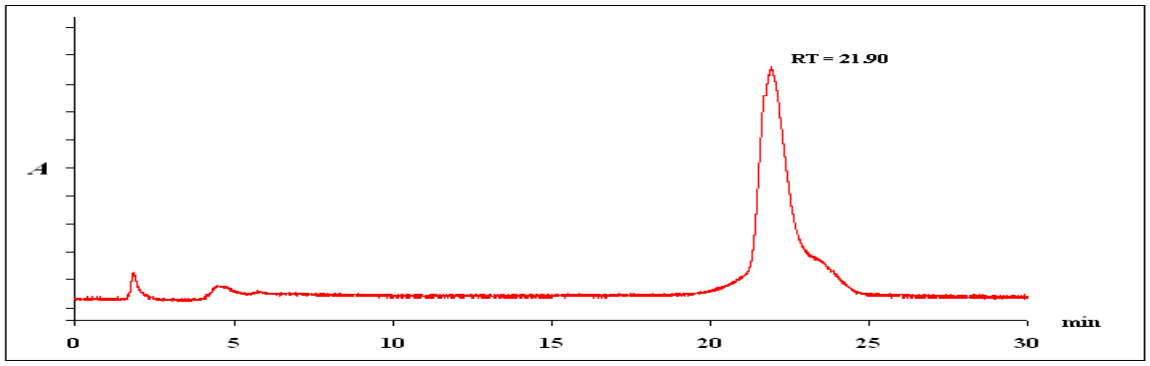


Figure 3.1 FPLC chromatogram of CG^{8CPh}

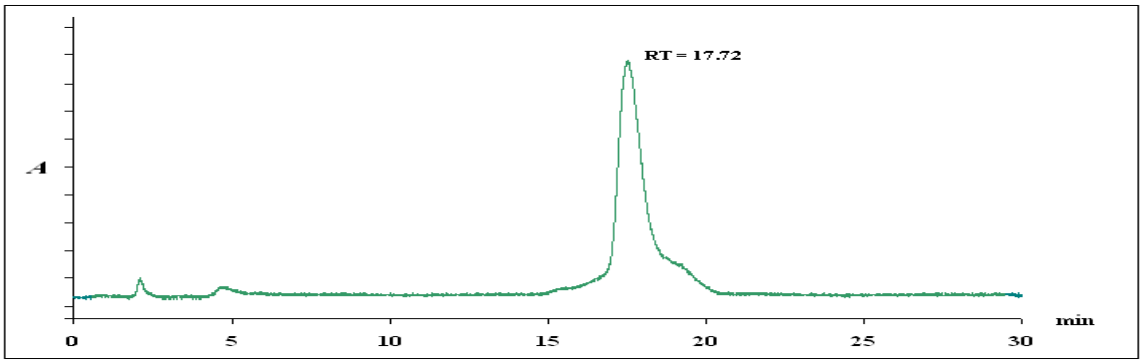


Figure 3.2 FPLC chromatogram of CG^{8MMPh}

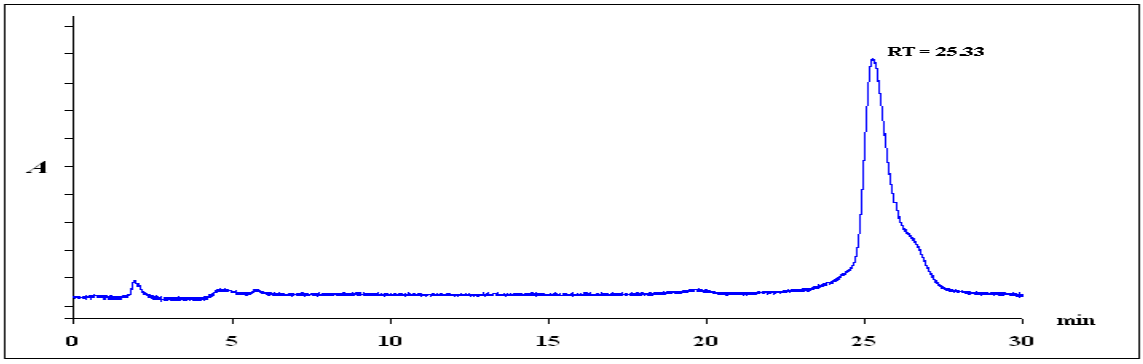


Figure 3.3 FPLC chromatogram of CG^{8HMPh}

3.2.4 ESI-MS Analysis

The purified CG decamers were examined by ESI-MS in negative ion mode to confirm that we have obtained the desired modified oligonucleotides based on their MW. The oligonucleotide sample (10 μM in DI water) was analyzed on a Finnigan LCQ deca using direct injection through a syringe pump. Due to the nature of oligonucleotides that have negatively charged phosphate backbones at physiological pH, the m/z of oligonucleotides are best detected when operating in the negative ion mode. The series of m/z peaks, based on charge state (e.g., M^{6-}, M^{5-}, M^{4-}, and M^{3-}) of the oligonucleotides with different number of sodium adducts are, generally observed. The observed m/z of CG^{8CPh}, CG^{8MMPh}, and CG^{8HMPh} were in agreement with the calculated m/z of the DNA as shown in Table 3.1. The result from ESI-MS indicates that the synthesized DNA have been successfully made and ready for structural and conformational analysis. The mass spectra of CG^{8CPh}, CG^{8MMPh}, and CG^{8HMPh} can be found in Appendix C.

Table 3.1 ESI-MS analysis for the synthesized oligonucleotides

DNA	MW (g/mol)	Calculated m/z	Found m/z
CG^{8CPh}	3150	524.0[a], 629.0[b], 786.5[c], 1049.0[d]	524.7, 629.8, 786.9, 1049.3
CG^{8MMPh}	3150	524.0, 629.0, 786.5, 1049.0	524.0, 629.2, 786.8, 1049.1
CG^{8HMPh}	3136	521.7, 626.2, 783.0, 1044.3	626.5, 783.4, 1044.5

[a] Charge state M^{6-}
[b] Charge state M^{5-}
[c] Charge state M^{4-}
[d] Charge state M^{3-}

The C8-carboxyphenyl, methoxymethylphenyl, and TBS protected hydroxymethylphenyl dG phosphoramidites have been made and characterized by NMR,

ESI-MS, and UV analysis. The Suzuki cross coupling was proven to be a convenient and efficient method to make C8-arylguanine adducts. The general protocol for making phosphoramidites that was developed for preparing the C8-phenyl and C8-tolyl modified dG seems to be generally applicable for other derivatives with minor alterations due to the small distinctions in physicochemical properties of the modified dG when different *para*-substituents were presented. For instance, the coupling between 8-BrdG and a relatively hydrophobic TBS protected 4-HMPBA required more content of ACN in Suzuki coupling solvent system compares to the syntheses of other aryl modified dG in this study. Use lower ACN content results in a precipitation of **(3)** and a very low production of the coupling product.

The N2 protection has been shown to be a straight forward step and gives excellent yields (70-90%) for all derivatives without the need for further purification, The 5′-OH protection step has usually resulted in lower yields and requires a purification step. Normal phase column chromatography (Al_2O_3, gradient 0-7% methanol in dichloromethane) were used to separate 5′-OH protection products with the exception of the C8-carboxyphenyl derivative, compound **(10)**, which was found to be difficult to separate with the chromatographic conditions successfully used with compounds **(11)** and **(12)**. Crystallization in dichloromethane seems to be a more practical alternative to purify compound **(10)**, although the product yield was low.

The phosphoramidites were made successfully and the modified DNA, CG^{8CPh}, CG^{8MMPh}, and CG^{8HMPh}, have been synthesized. As with phosphoramidites of the C8-phenyl and tolyl, the phosphoramidites prepared here form gels shortly after dissolving in ACN. This property necessitates manual addition of the modified base. We do not know

the exact nature of the gels formed though G-quartet formation is a reasonable possibilitity. After FPLC purification, ESI-MS was used to identify the oligonucleotides based on their MW. The result from ESI-MS indicates the success of CG^{8CPh}, CG^{8MMPh}, and CG^{8HMPh} syntheses.

CHAPTER 4
EXPERIMENTAL PROCEDURES AND RESULTS PART II: CD AND NMR ANALYSES OF CG DECAMERS CONTAINING C8-ARYL MODIFIED GUANINE

In previous work we demonstrated that the incorporation of a C8-phenylguanine adduct in a CG sequence stabilizes the Z DNA form and/or destabilizes the B-DNA form[22]. Additional C8-arylguanine adducts had also been examined by molecular modeling and free energy calculations[99]. The results of the computational methods were in agreement with the experimental results that were available (e.g., unmodified and C8-phenyl adduct). However, the computational results and predictions for some aryl hydrazine related C8-adducts were not consistent with the proposed theory that Z DNA is involved in aryl hydrazine carcinogenesis (e.g., C8-hydroxymethylphenyl adduct). Therefore, to determine whether the results from prior computational studies, were correct, the CG decamers, CG^{8Tol}, CG^{8CPh}, CG^{8MMPh} and CG^{8HMPh}, were prepared and analyzed by CD and NMR

Circular dichroism is a useful tool for the study of B and Z DNA as the spectra can be used to demonstrate the presence of these forms of DNA both qualitatively and quantitatively. Here, we have used CD data to determine the molar fractions of B-DNA, Z-DNA, and ssDNA of C8-arylguanine modified oligonucleotides as a function of temperature and salt concentration. These data permit the comparison of the effects of the selected aryl adducts on Z-DNA stabilization/B-DNA destabilization in a quantitative fashion and provides insight in the nature and extent of the Z-DNA stabilization/B-DNA destabilization aryl adducts produce. Ultimately, the goal is to determine, how the aryl

adducts facilitate B-Z conversion, how they stabilize the Z DNA form and/or how they destabilize the B form. In addition, the carcinogenic profiles and B-Z-DNA transition effects of the selected aryl hydrazines can be correlated and evaluated, hence the potential role of the Z form in carcinogenesis can be assessed.

All CG sequences used in this study have been examined with CD and NMR analysis. The conformation of CG decamers can be assumed from CD spectra, while the B-Z conversion can be observed by the change in pattern of the spectra. The proton NMR experiments provide useful pieces of information regarding DNA conformation that have proven to be useful to identify DNA conformation and support CD data.

4.1 CD Analysis and Molar Fraction Calculation of CG Decamers Containing C8-Aryl Modified Guanine

4.1.1 General

The oligonucleotide samples (50 μM of CG, CG^{8Ph}, CG^{8Tol}, CG^{8CPh}, CG^{8MMPh} or CG^{8HMPh} in 10 mM phosphate buffer, pH 7.4, 300 μL) with various sodium chloride concentrations (0, 25, 50, 100, 200, 500, 1000, 2000, and 4000 mM) were prepared in a centrifuge tube. Prior to acquiring CD spectra, oligonucleotide samples were annealed by heating at 90°C for 30 min, slowly cooling down to room temperature, and then further cooled to 4°C for and additional 15 min. A sample was transferred to cuvette and CD spectra were obtained using a Jasco J-810 spectropolarimeter (Easton, MD).

From the CD results, the molar fractions of B, Z and ssDNA, over the temperature range 10-90°C and with NaCl concentrations from 0-4000 mM, were calculated as described previously in Section 2.4.2

4.1.2 CD Measurment

The CD spectra of each modified CG decamer sample was recorded from 220 to 350 nm at 10, 20, 30, 37, 50, 60, 70, 80, and 90°C using the instruments temperature/wavelength scan mode with the temperature ramp rate of 1°C/min and a scan rate of 50 nm/min. The data reported are the averages of two duplicate experiments.

The tertiary structure of DNA such as B and Z-DNA can be distinguished by CD. The CD signal of B DNA shows positive ellipticity at approximately 280 nm and 220 nm and a negative ellipticity at approximately 250 nm. In contrast, the CD spectrum Z-DNA shows negative ellipticity at approximately 295 nm. The CD spectra of CG and aryl modified CG, as a function of NaCl concentrations from 0-4000 mM at 37°C, are shown in Figure 4.1-4.6. Generally, at low salt concentration, the alternating CG sequences adopt a B-DNA conformation, while under high salt conditions a conformational conversion for B to Z can be observed. The conversion of B to Z-DNA can be monitored by the change in CD spectra. The negative ellipticipy observed at 295 nm increases in intensity while, simultaneously, the negative ellipticity at 250 nm decreases in intensity as the B-DNA converts to Z-DNA.

In agreement with previous studies of CG oligomers[82], the unmodified CG decamer predominantly adopts the B conformation at salt concentrations below 2 M, as indicated by the positive ellipticity at 280 nm and negative ellipticity at approximately 250 nm (Figure 4.1). At 2 M NaCl, there is only a small amount of the Z form observed and as the salt concentration is further increase to 4 M, the Z-DNA becomes predominant (i.e., the positive ellipticity at 280 nm is replaced by negative ellipticity peak at 295 nm).

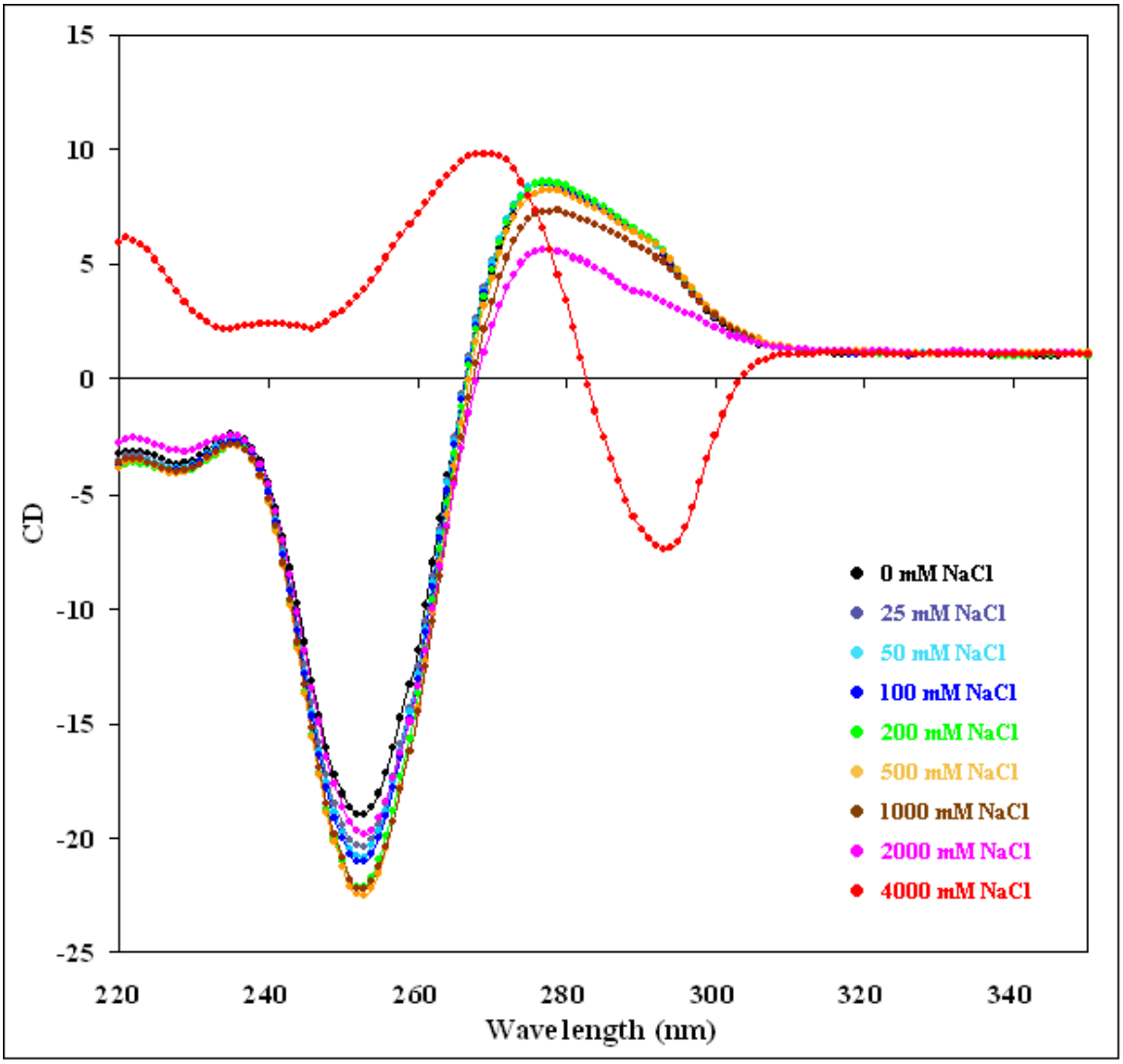


Figure 4.1 CD spectra of CG with NaCl 0-4000 mM at 37°C

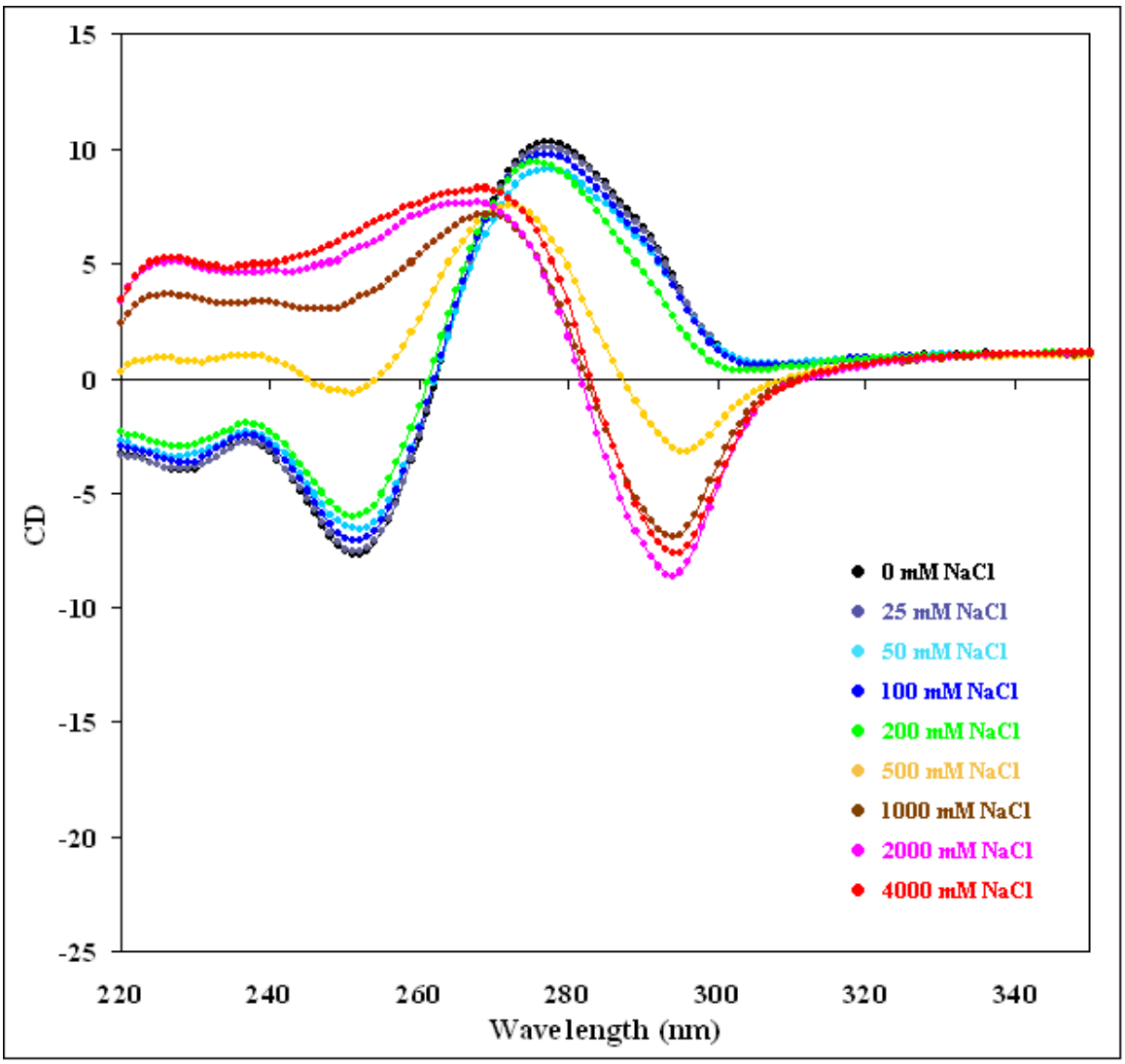


Figure 4.2 CD spectra of CG^{8Ph} with NaCl 0-4000 mM at 37°C

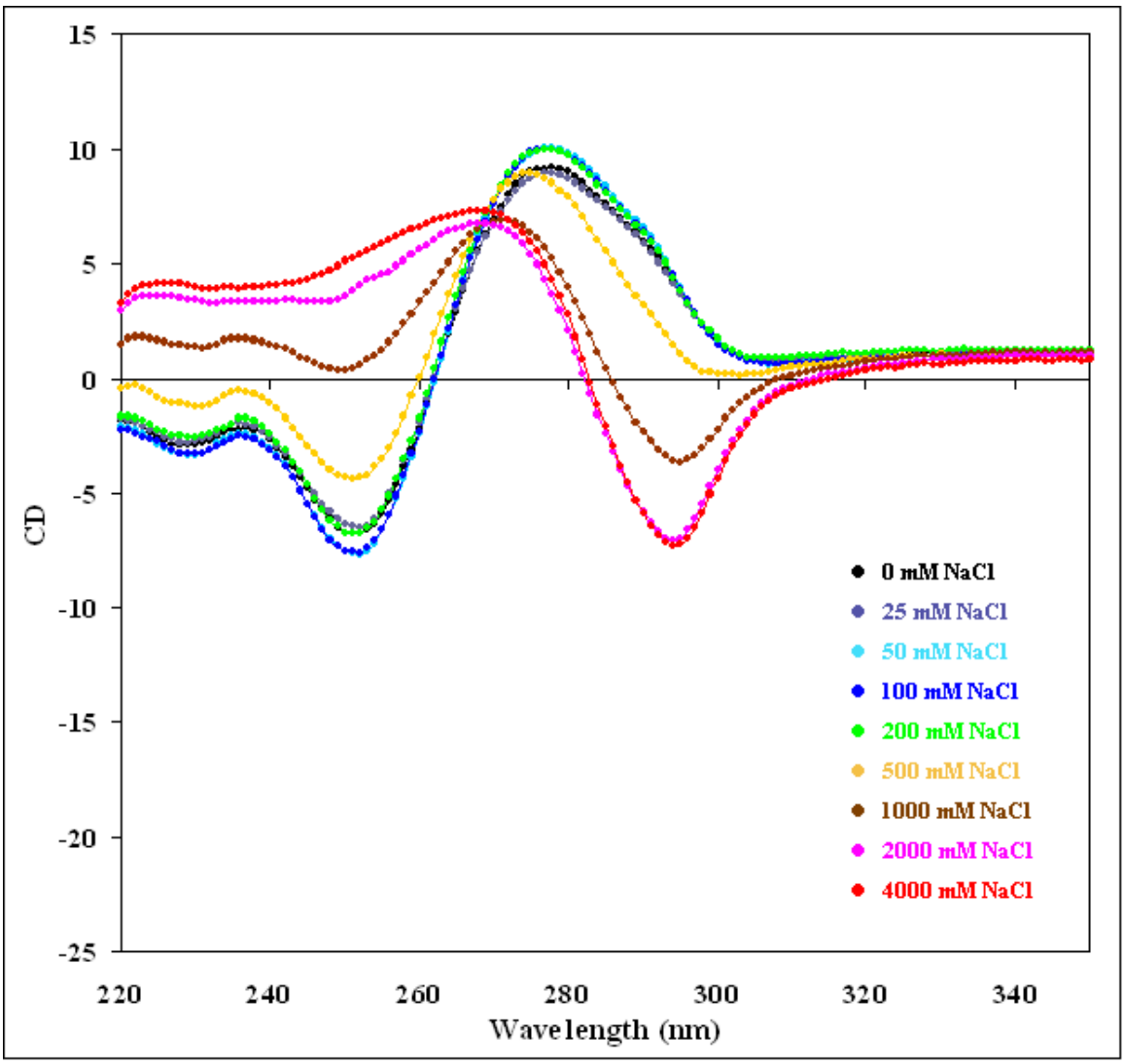


Figure 4.3 CD spectra of CG^{8Tol} with NaCl 0-4000 mM at 37°C

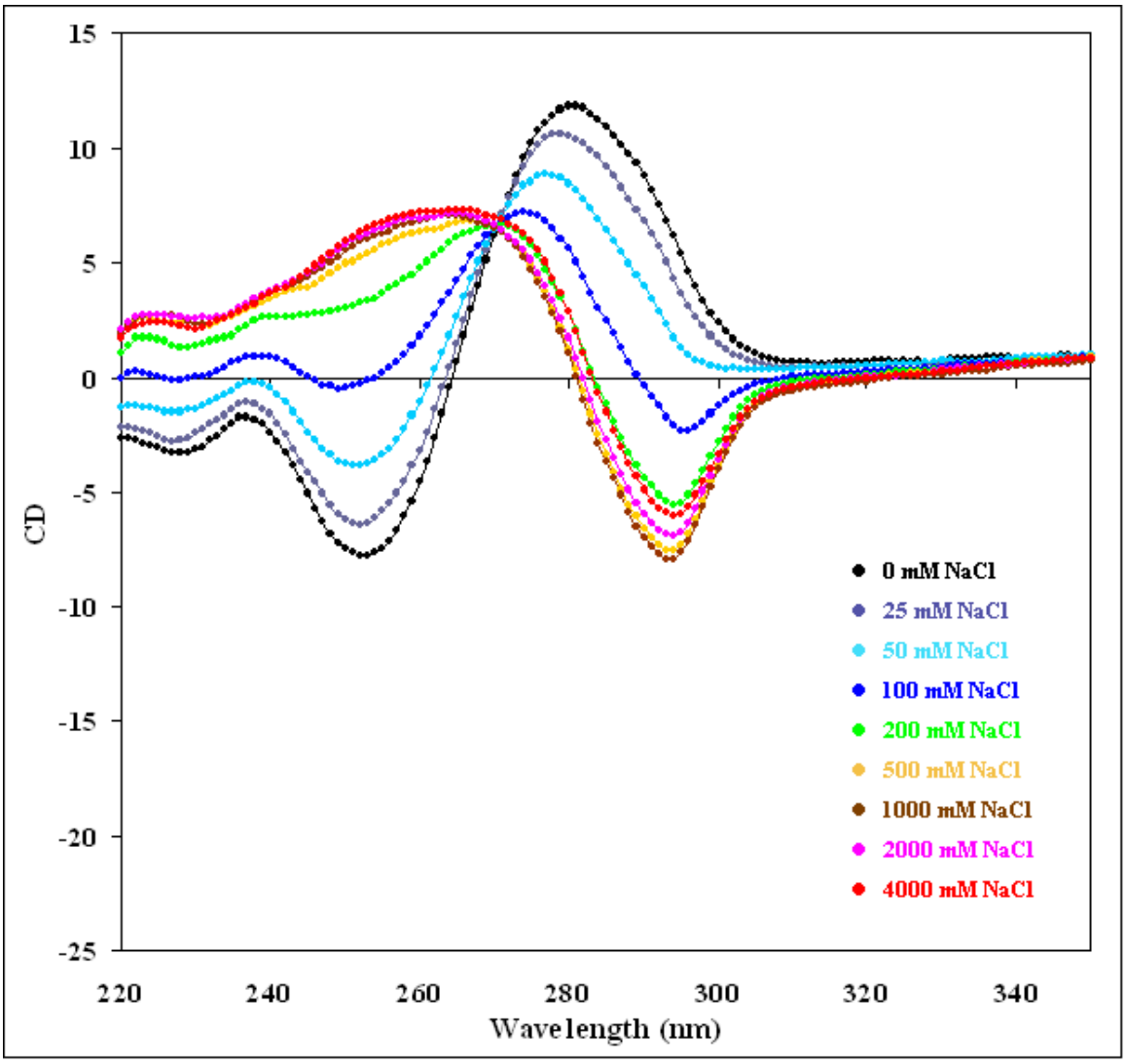


Figure 4.4 CD spectra of CG^{8CPh} with NaCl 0-4000 mM at 37°C

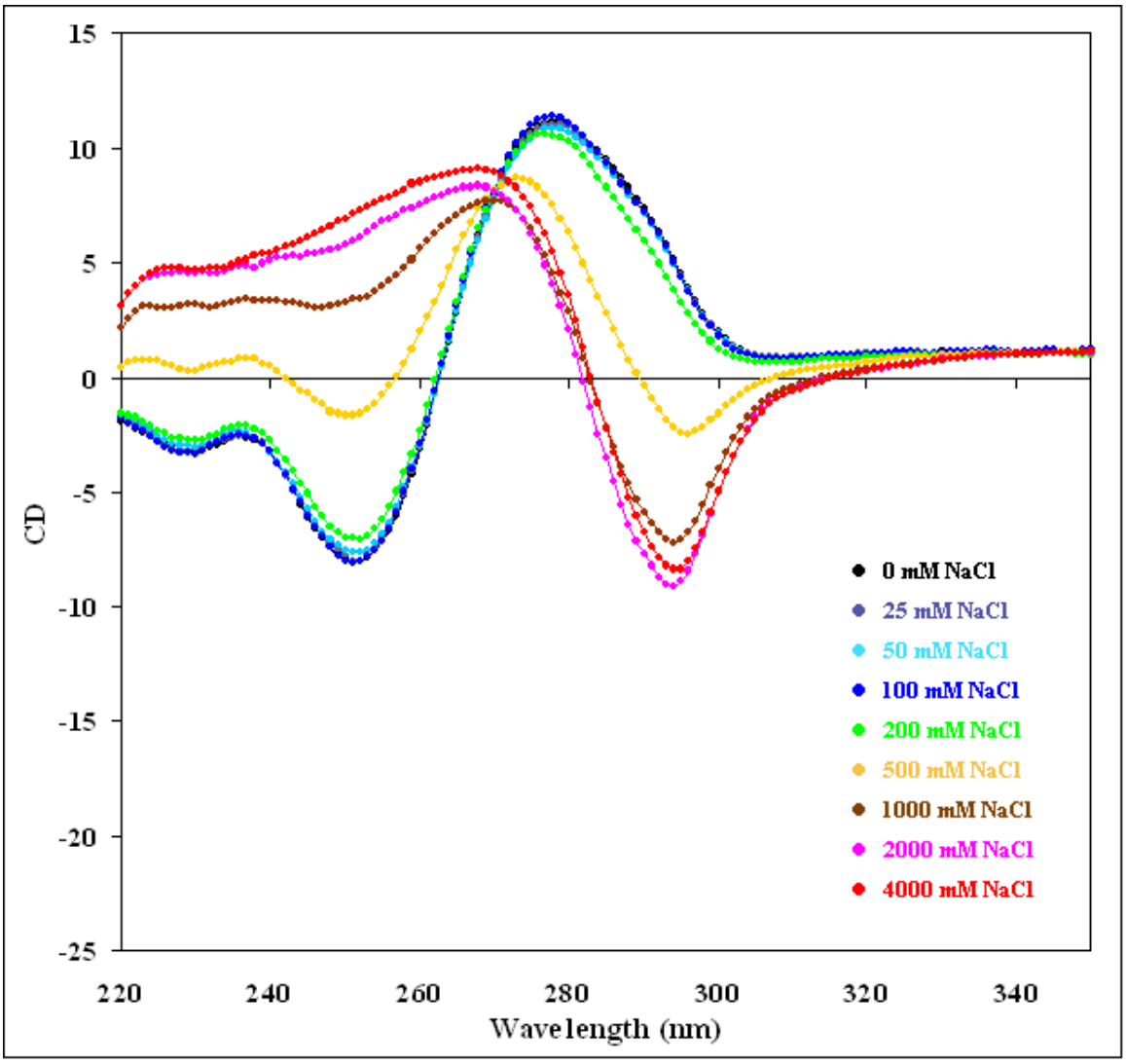


Figure 4.5 CD spectra of CG^{8MMPh} with NaCl 0-4000 mM at 37°C

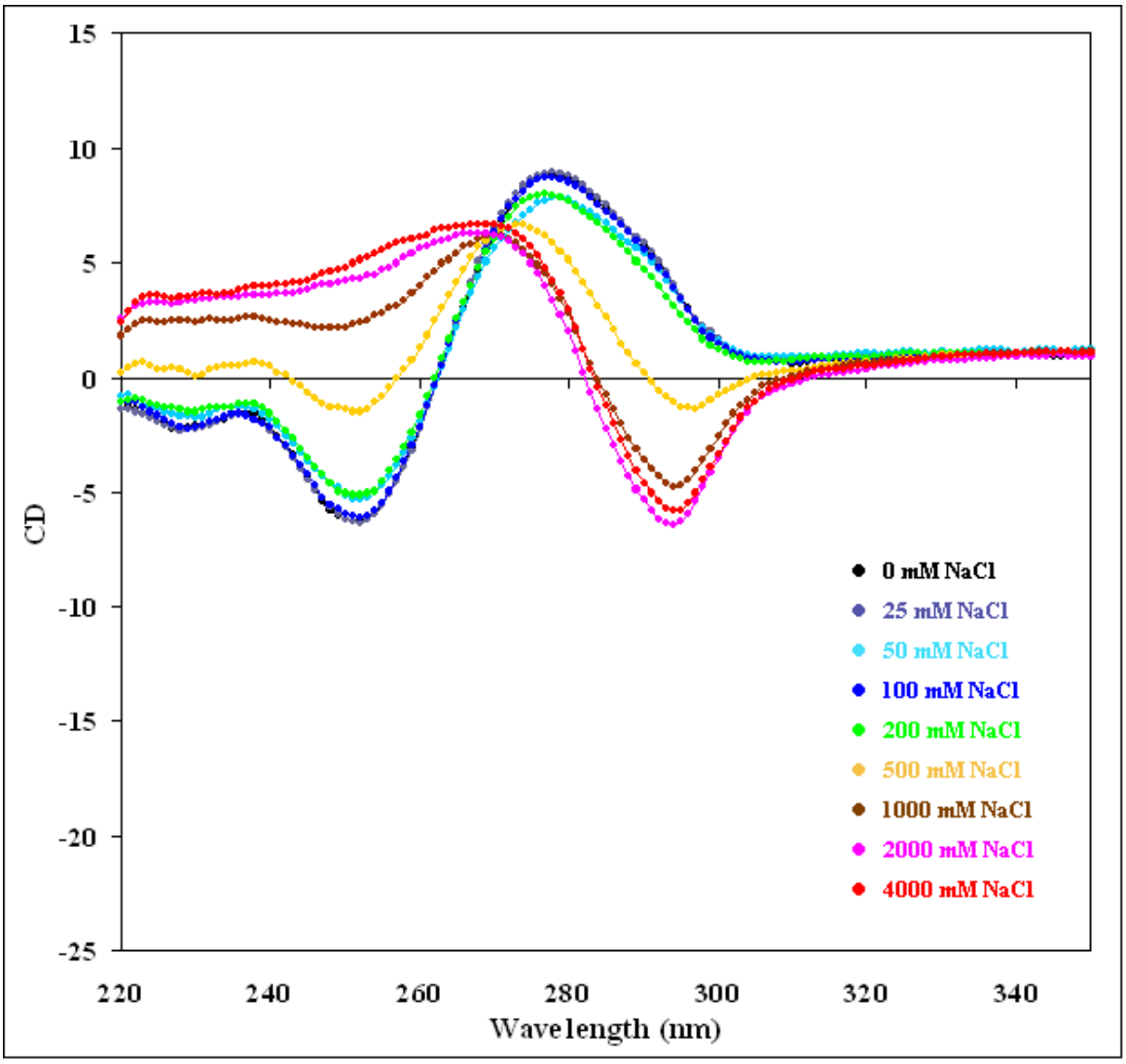


Figure 4.6 CD spectra of CG^{8HMPh} with NaCl 0-4000 mM at 37°C

The CD spectra of CG^{8Ph} (Figure 4.2), CG^{8MMPh} (Figure 4.5), and CG^{8HMPh} (Figure 4.6), show that the Z form is predominant (B-Z conversion greater than 50%) in samples that have salt concentration of approximately 500 mM, considerably less than the 2 M salt concentration required for the unmodified oligonucleotide. The modified CG decamers are approximately 100% in the Z conformation at 1 M salt concentrations as indicated by CD spectra. Based on the lower concentration of sodium chloride needed to stabilize the Z comformation of CG^{8Ph}, CG^{8MMPh}, and CG^{8HMPh} as compared to the unmodified CG, indicates Z-DNA stabilization and/or B-DNA destabilization effects of C8-phenyl, methoxymethylphenyl, hydroxymethylphenyl guanine adducts.

The effect of C8-tolyl guanine adduct on Z-DNA formation seem to be lower since CG^{8Tol} (Figure 4.3) requires 500 mM to 1 M salt to facilitate B-Z conversion and requires approximately 2 M NaCl to make the Z-DNA form predominant. On the other hand, the C8-carboxyphenyl guanine adduct seems to have the greatest effect on Z-DNA formation, due to the fact that the B-Z conversion of CG^{8CPh} (Figure 4.4) is observed at very low salt concentration (25- 50 mM) and the oligonucleotide requires only 200 mM salt to entirely adopt in Z conformation.

Generally, the C8-arylguanine modified CG decamers principally adopt the Z form at salt concentrations below 1 M, much lower that what is needed for the unmodified CG, indicating the Z-DNA stabilization and/or B-DNA destabilization effects of the aryl adducts on CG decamer sequences. Among the aryl adducts, *p*-carboxyphenyl has the greatest effect while tolyl has the lowest compare with phenyl, *p*-methoxymethylphenyl, and *p*-hydroxymethylphenyl which have comparable effects on Z-DNA stabilization. To compare the effect of each aryl adduct on Z-DNA formation, the

molar fraction of Z-DNA in each sample has been determined as described in the next part.

4.1.3 Molar Fraction Calculation

CD spectra were analyzed assuming that there is equilibrium among three conformation of DNA (B, Z, and ssDNA) in solution. Therefore, at each temperature

$$\Delta\epsilon^{295} = \Delta\epsilon_B^{295} * f_B + \Delta\epsilon_Z^{295} * f_Z + \Delta\epsilon_{ss}^{295} * f_{ss} \qquad \text{Equation 3.1}$$

$$1 = f_B + f_Z + f_{ss} \qquad \text{Equation 3.2}$$

While f is the molar fraction of each DNA conformation, $\Delta\epsilon$ of each conformation was estimated from CD at 295 nm. At 270 nm, the isosbestic point of dsDNA forms and where $\Delta\varepsilon$ of B and Z are equal, f_{ss} was determined from the Equations 3.3 and 3.4. The f_B and f_Z then were solved from the calculated f_{ss}, Equation 3.1, and 3.2.

$$\Delta\varepsilon^{270} = \Delta\varepsilon_{ds}^{270} * f_{ds} + \Delta\varepsilon_{ss}^{270} * f_{ss} \qquad \text{Equation 3.3}$$

$$1 = f_{ds} + f_{ss} \qquad \text{Equation 3.4}$$

Generally, for each oligonucleotide, $\Delta\varepsilon_{ds}^{270}$ was obtained and averaged from UV absorbance at 270 nm from all samples at 10°C where DNA denaturation is minimized, while $\Delta\varepsilon_{ss}^{270}$ was obtained the same way but at 90°C assuming that all DNA are in single strand form. Molar ellipticities of B-DNA, Z-DNA and ssDNA are estimated from CD at 295 nm under conditions such that only one conformation of DNA is predominant. At 30°C, without or with low salt $\Delta\epsilon_B^{295}$, was obtained, while at 10°C with high salt

concentration $\Delta\epsilon_Z^{295}$ was calculated. The $\Delta\epsilon_{ss}^{295}$ was estimated from CD^{295} of the DNA samples at 90°C.

The f_B, f_Z, and f_{ss} of CG, CG^{8Ph}, CG^{8Tol}, CG^{8CPh}, CG^{8MMPh} and CG^{8HMPh} were approximated and are shown in Tables 4.1 through 4.6, respectively. From the calculated data, the conformation diagrams, that represent the molar fractions of each DNA conformation as a function of temperatures and sodium chloride concentration, have been plotted and are shown in Figures 4.7- 4.12 for CG, CG^{8Ph}, CG^{8Tol}, CG^{8CPh}, CG^{8MMPh} and CG^{8HMPh}, respectively. The diagrams are consistent with the known behavior of DNA where B-DNA predominates at low salt concentration or higher temperatures and Z-DNA formation is favored at lower temperatures and higher salt concentrations. Raising the temperature too high, of course, causes DNA denaturation and results in the formation of ssDNA.

Table 4.1 Molar fractions of B-DNA, Z-DNA, and ssDNA of CG with salt 0-4000 mM at temperature 10-90°C

NaCl (mM)	f_B at Temperature (°C)								
	10	20	30	37	50	60	70	80	90
0	0.99	0.96	0.93	0.90	0.74	0.52	0.24	0.12	0.00
25	0.98	0.99	0.99	0.98	0.92	0.78	0.44	0.16	0.00
50	0.98	0.98	0.98	0.96	0.92	0.83	0.54	0.20	0.00
100	0.99	0.97	0.95	0.94	0.90	0.83	0.60	0.22	0.00
200	0.98	0.99	0.99	0.98	0.96	0.90	0.69	0.28	0.00
500	0.97	0.98	0.99	0.99	0.96	0.91	0.73	0.34	0.00
1000	0.96	0.96	0.96	0.95	0.91	0.84	0.67	0.31	0.00
2000	0.72	0.80	0.86	0.88	0.88	0.79	0.64	0.24	0.00
4000	0.00	0.07	0.16	0.22	0.34	0.40	0.34	0.13	0.00

NaCl (mM)	f_Z at Temperature (°C)								
	10	20	30	37	50	60	70	80	90
0	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
25	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
50	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
100	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
200	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
500	0.03	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00
1000	0.04	0.03	0.02	0.01	0.00	0.00	0.00	0.00	0.00
2000	0.28	0.18	0.12	0.09	0.05	0.01	0.00	0.00	0.00
4000	1.00	0.89	0.77	0.68	0.51	0.36	0.17	0.03	0.00

NaCl (mM)	f_{ss} at Temperature (°C)								
	10	20	30	37	50	60	70	80	90
0		0.04	0.07	0.10	0.26	0.48	0.76	0.88	1.00
25	0.00	0.00	0.01	0.02	0.08	0.22	0.56	0.84	1.00
50	0.00	0.00	0.02	0.04	0.08	0.17	0.46	0.80	1.00
100	0.00	0.03	0.05	0.06	0.10	0.17	0.40	0.78	1.00
200	0.00	0.00	0.01	0.02	0.04	0.10	0.31	0.72	1.00
500	0.00	0.00	0.00	0.01	0.04	0.09	0.27	0.66	1.00
1000	0.00	0.01	0.03	0.04	0.09	0.16	0.33	0.69	1.00
2000	0.00	0.02	0.02	0.03	0.08	0.20	0.36	0.76	1.00
4000	0.00	0.04	0.07	0.10	0.15	0.24	0.49	0.84	1.00

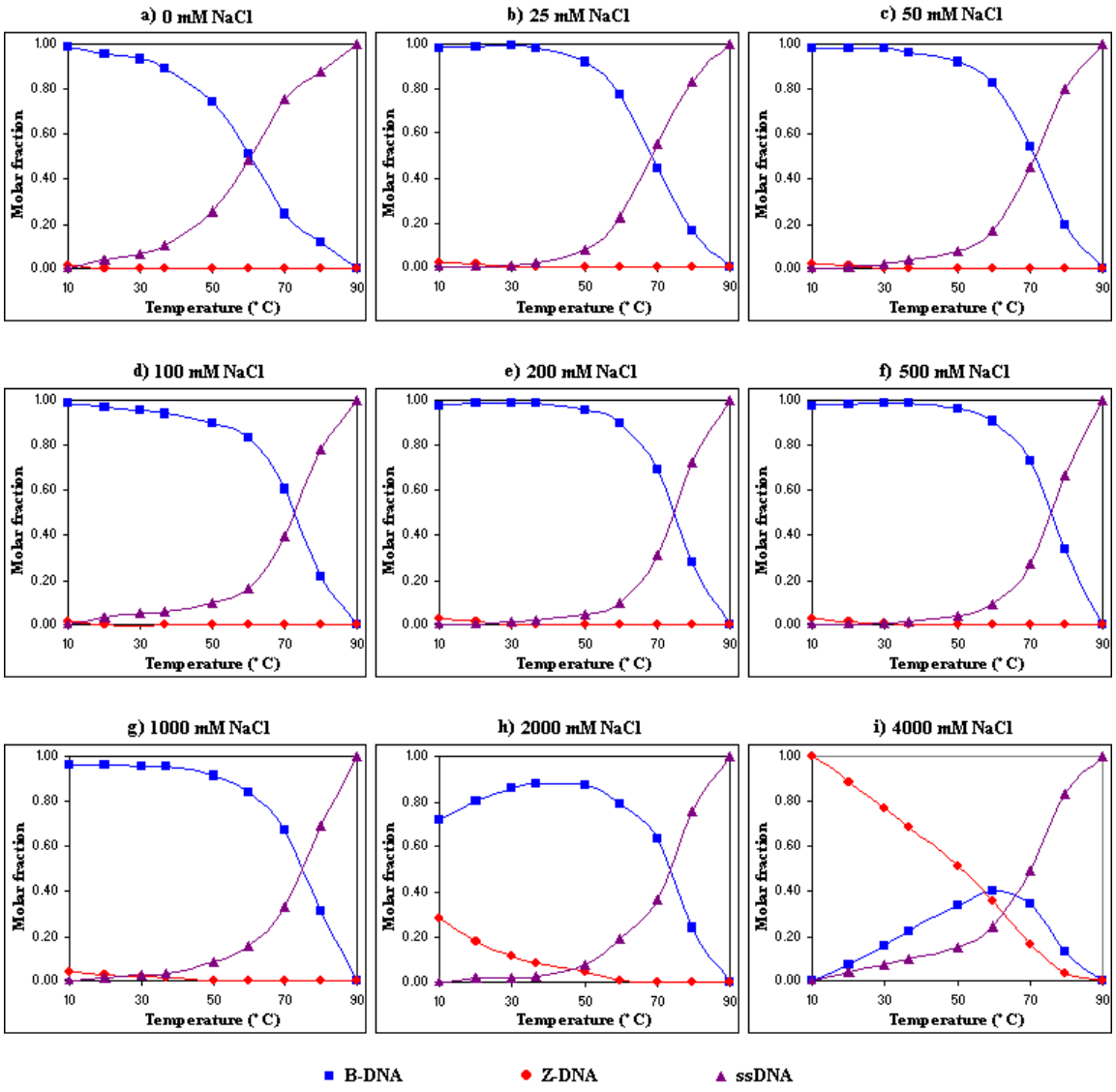


Figure 4.7 Plots of the molar fractions of B-DNA, Z-DNA, and ssDNA of CG over the temperature range 10-90°C. The oligonucleotide samples were in 10 mM sodium phosphate buffer, pH 7.4 and 0 mM (a), 25 mM (b), 50 mM (c), 100 mM (d), 200 mM (e), 500 mM (f), 1000 mM (g), 2000 mM (h), and 4000 mM (i) NaCl.

Table 4.2 Molar fractions of B-DNA, Z-DNA, and ssDNA of CG^{8Ph} with salt 0-4000 mM at temperature 10-90°C

NaCl (mM)	f_B at Temperature (°C)								
	10	20	30	37	50	60	70	80	90
0	0.98	0.96	0.96	0.96	0.92	0.79	0.54	0.25	0.00
25	0.77	0.85	0.89	0.88	0.85	0.75	0.54	0.27	0.00
50	0.85	0.90	0.97	0.96	0.94	0.84	0.61	0.30	0.00
100	0.65	0.68	0.85	0.88	0.83	0.73	0.53	0.25	0.00
200	0.32	0.36	0.65	0.77	0.76	0.69	0.48	0.23	0.00
500	0.01	0.06	0.24	0.38	0.59	0.58	0.42	0.21	0.00
1000	0.00	0.01	0.10	0.19	0.41	0.49	0.38	0.17	0.00
2000	0.00	0.04	0.09	0.14	0.25	0.35	0.29	0.12	0.00
4000	0.04	0.11	0.16	0.19	0.24	0.26	0.20	0.08	0.00

NaCl (mM)	f_Z at Temperature (°C)								
	10	20	30	37	50	60	70	80	90
0	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
25	0.23	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00
50	0.15	0.10	0.01	0.00	0.00	0.00	0.00	0.00	0.00
100	0.35	0.32	0.05	0.00	0.00	0.00	0.00	0.00	0.00
200	0.68	0.64	0.24	0.08	0.00	0.00	0.00	0.00	0.00
500	0.99	0.89	0.63	0.42	0.07	0.00	0.00	0.00	0.00
1000	1.00	0.98	0.85	0.72	0.35	0.09	0.00	0.00	0.00
2000	1.00	0.96	0.90	0.84	0.65	0.38	0.10	0.02	0.00
4000	0.96	0.89	0.83	0.79	0.69	0.54	0.27	0.06	0.00

NaCl (mM)	f_{ss} at Temperature (°C)								
	10	20	30	37	50	60	70	80	90
0	0.00	0.04	0.04	0.04	0.08	0.21	0.46	0.75	1.00
25	0.00	0.10	0.11	0.12	0.15	0.25	0.46	0.73	1.00
50	0.00	0.00	0.02	0.04	0.06	0.16	0.39	0.69	1.00
100	0.00	0.01	0.11	0.12	0.17	0.27	0.47	0.75	1.00
200	0.00	0.00	0.10	0.15	0.24	0.31	0.52	0.77	1.00
500	0.00	0.05	0.13	0.20	0.33	0.42	0.58	0.79	1.00
1000	0.00	0.01	0.04	0.08	0.24	0.42	0.62	0.82	1.00
2000	0.00	0.00	0.01	0.02	0.11	0.27	0.61	0.86	1.00
4000	0.00	0.00	0.02	0.02	0.07	0.20	0.53	0.86	1.00

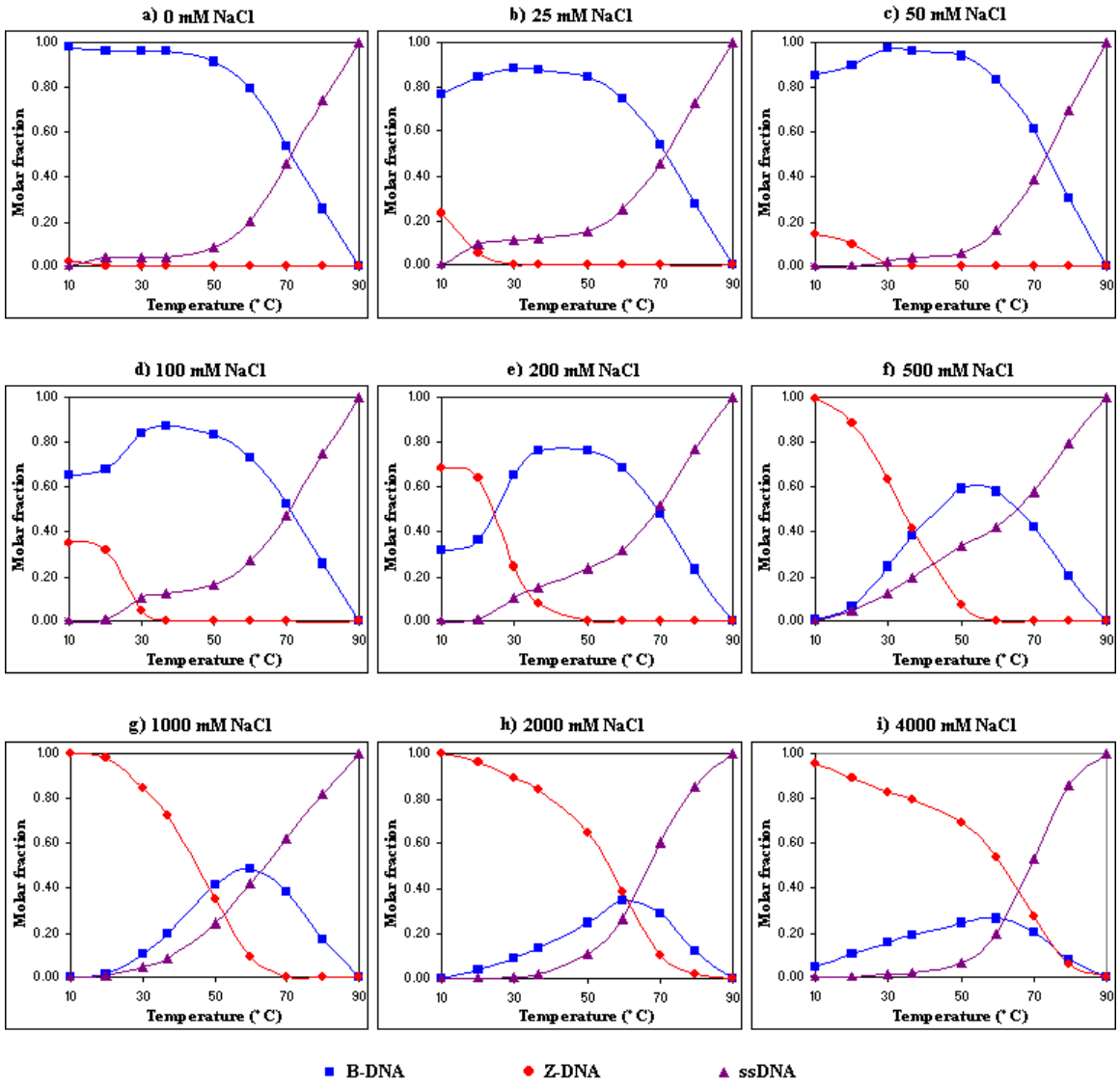


Figure 4.8 Plots of the molar fractions of B-DNA, Z-DNA, and ssDNA of CG^{8Ph} over the temperature range 10-90°C. The oligonucleotide samples were in 10 mM sodium phosphate buffer, pH 7.4 and 0 mM (a), 25 mM (b), 50 mM (c), 100 mM (d), 200 mM (e), 500 mM (f), 1000 mM (g), 2000 mM (h), and 4000 mM (i) NaCl.

Table 4.3 Molar fractions of B-DNA, Z-DNA, and ssDNA of CG^{8Tol} with salt 0-4000 mM at temperature 10-90°C

NaCl (mM)	f_B at Temperature (°C)								
	10	20	30	37	50	60	70	80	90
0	0.98	0.95	0.94	0.93	0.81	0.67	0.42	0.22	0.00
25	0.94	0.97	0.92	0.81	0.62	0.43	0.34	0.24	0.00
50	0.71	0.84	0.87	0.89	0.86	0.78	0.58	0.30	0.00
100	0.37	0.67	0.76	0.75	0.66	0.70	0.53	0.27	0.00
200	0.16	0.46	0.71	0.75	0.75	0.68	0.53	0.26	0.00
500	0.04	0.18	0.44	0.59	0.67	0.61	0.46	0.22	0.00
1000	0.00	0.03	0.20	0.34	0.55	0.56	0.42	0.20	0.00
2000	0.00	0.02	0.09	0.16	0.33	0.41	0.33	0.15	0.00
4000	0.00	0.04	0.10	0.14	0.20	0.24	0.20	0.09	0.00

NaCl (mM)	f_Z at Temperature (°C)								
	10	20	30	37	50	60	70	80	90
0	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
25	0.06	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
50	0.29	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00
100	0.63	0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00
200	0.84	0.38	0.04	0.00	0.00	0.00	0.00	0.00	0.00
500	0.96	0.75	0.36	0.14	0.00	0.00	0.00	0.00	0.00
1000	1.00	0.93	0.70	0.49	0.12	0.00	0.00	0.00	0.00
2000	1.00	0.96	0.86	0.76	0.47	0.19	0.03	0.02	0.00
4000	1.00	0.94	0.87	0.82	0.67	0.46	0.18	0.05	0.00

NaCl (mM)	f_{ss} at Temperature (°C)								
	10	20	30	37	50	60	70	80	90
0	0.00	0.05	0.06	0.07	0.19	0.33	0.58	0.78	1.00
25	0.00	0.01	0.08	0.19	0.38	0.57	0.66	0.76	1.00
50	0.00	0.12	0.13	0.11	0.14	0.22	0.42	0.70	1.00
100	0.00	0.17	0.24	0.25	0.34	0.30	0.47	0.73	1.00
200	0.00	0.16	0.26	0.25	0.25	0.32	0.47	0.74	1.00
500	0.00	0.07	0.20	0.27	0.33	0.39	0.54	0.78	1.00
1000	0.00	0.04	0.10	0.17	0.33	0.44	0.58	0.80	1.00
2000	0.00	0.03	0.05	0.08	0.20	0.41	0.64	0.83	1.00
4000	0.00	0.02	0.02	0.04	0.13	0.30	0.62	0.86	1.00

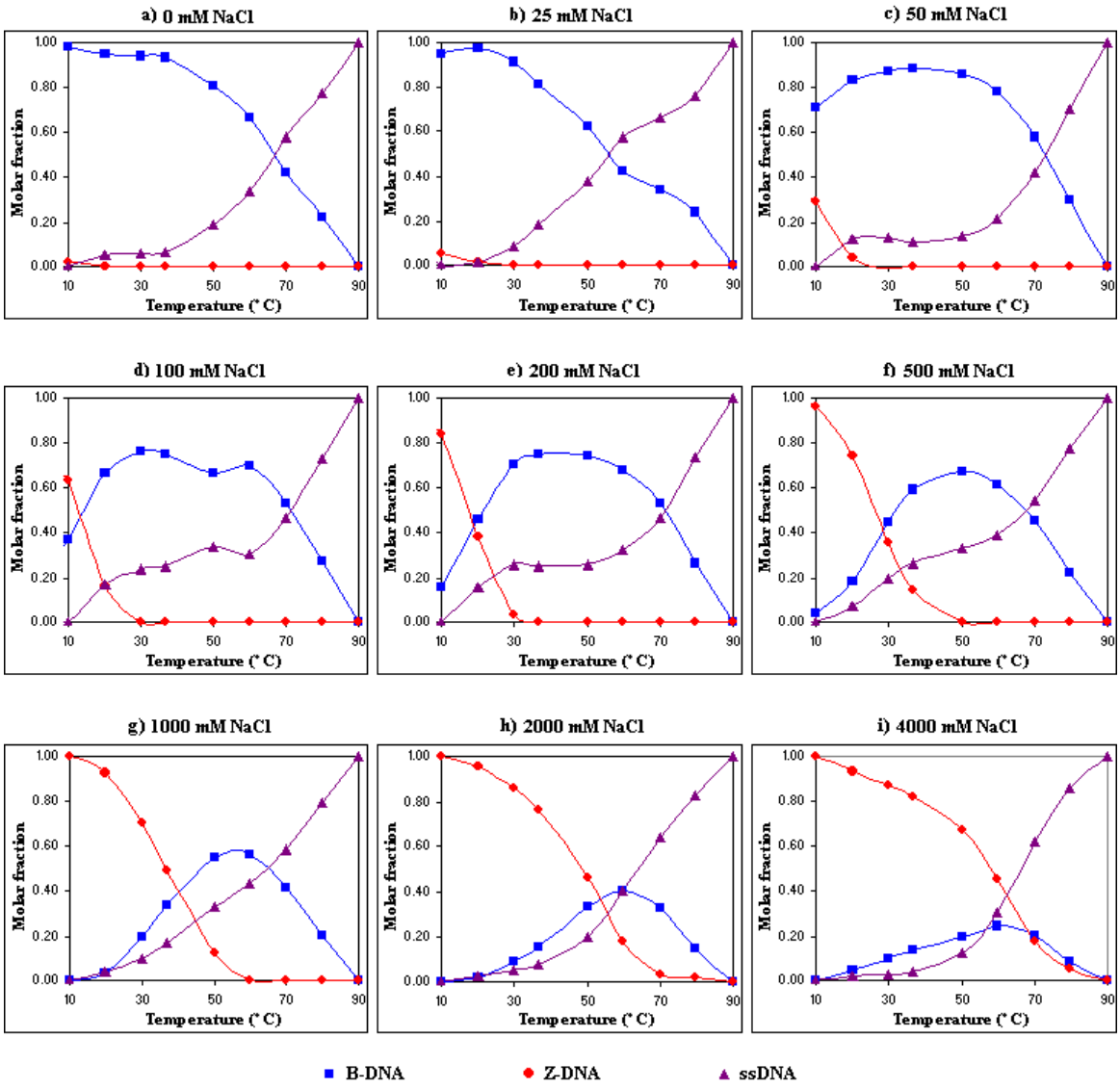


Figure 4.9 Plots of the molar fractions of B-DNA, Z-DNA, and ssDNA of CG^{8Tol} over the temperature range 10-90°C. The oligonucleotide samples were in 10 mM sodium phosphate buffer, pH 7.4 and 0 mM (a), 25 mM (b), 50 mM (c), 100 mM (d), 200 mM (e), 500 mM (f), 1000 mM (g), 2000 mM (h), and 4000 mM (i) NaCl.

Table 4.4 Molar fractions of B-DNA, Z-DNA, and ssDNA of CG^{8CPh} with salt 0-4000 mM at temperature 10-90°C

NaCl (mM)	f_B at Temperature (°C)								
	10	20	30	37	50	60	70	80	90
0	0.32	0.45	0.71	0.71	0.61	0.44	0.23	0.08	0.00
25	0.24	0.31	0.62	0.70	0.61	0.47	0.25	0.09	0.00
50	0.11	0.16	0.39	0.53	0.56	0.44	0.26	0.11	0.00
100	0.00	0.05	0.21	0.35	0.51	0.42	0.25	0.11	0.00
200	0.01	0.04	0.12	0.20	0.40	0.44	0.30	0.13	0.00
500	0.00	0.02	0.06	0.10	0.22	0.31	0.27	0.13	0.00
1000	0.00	0.02	0.05	0.08	0.14	0.21	0.21	0.12	0.00
2000	0.06	0.07	0.09	0.11	0.15	0.17	0.16	0.09	0.00
4000	0.09	0.12	0.14	0.16	0.18	0.19	0.16	0.07	0.00

NaCl (mM)	f_Z at Temperature (°C)								
	10	20	30	37	50	60	70	80	90
0	0.68	0.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00
25	0.76	0.69	0.22	0.05	0.00	0.00	0.00	0.01	0.00
50	0.89	0.79	0.39	0.16	0.00	0.00	0.00	0.00	0.00
100	1.00	0.89	0.61	0.39	0.04	0.00	0.00	0.00	0.00
200	0.99	0.92	0.77	0.63	0.24	0.03	0.00	0.00	0.00
500	1.00	0.94	0.85	0.79	0.54	0.25	0.04	0.01	0.00
1000	1.00	0.94	0.86	0.81	0.65	0.42	0.14	0.03	0.00
2000	0.94	0.87	0.81	0.77	0.66	0.51	0.25	0.05	0.00
4000	0.91	0.83	0.76	0.71	0.63	0.54	0.35	0.10	0.00

NaCl (mM)	f_{ss} at Temperature (°C)								
	10	20	30	37	50	60	70	80	90
0	0.00	0.12	0.29	0.29	0.39	0.56	0.77	0.92	1.00
25	0.00	0.00	0.16	0.25	0.39	0.53	0.75	0.90	1.00
50	0.00	0.06	0.22	0.31	0.44	0.56	0.74	0.89	1.00
100	0.00	0.07	0.18	0.27	0.46	0.58	0.75	0.89	1.00
200	0.00	0.04	0.11	0.17	0.35	0.53	0.70	0.86	1.00
500	0.00	0.04	0.09	0.12	0.24	0.45	0.69	0.85	1.00
1000	0.00	0.05	0.09	0.11	0.20	0.37	0.65	0.85	1.00
2000	0.00	0.06	0.11	0.12	0.19	0.32	0.59	0.86	1.00
4000	0.00	0.05	0.10	0.12	0.19	0.27	0.49	0.82	1.00

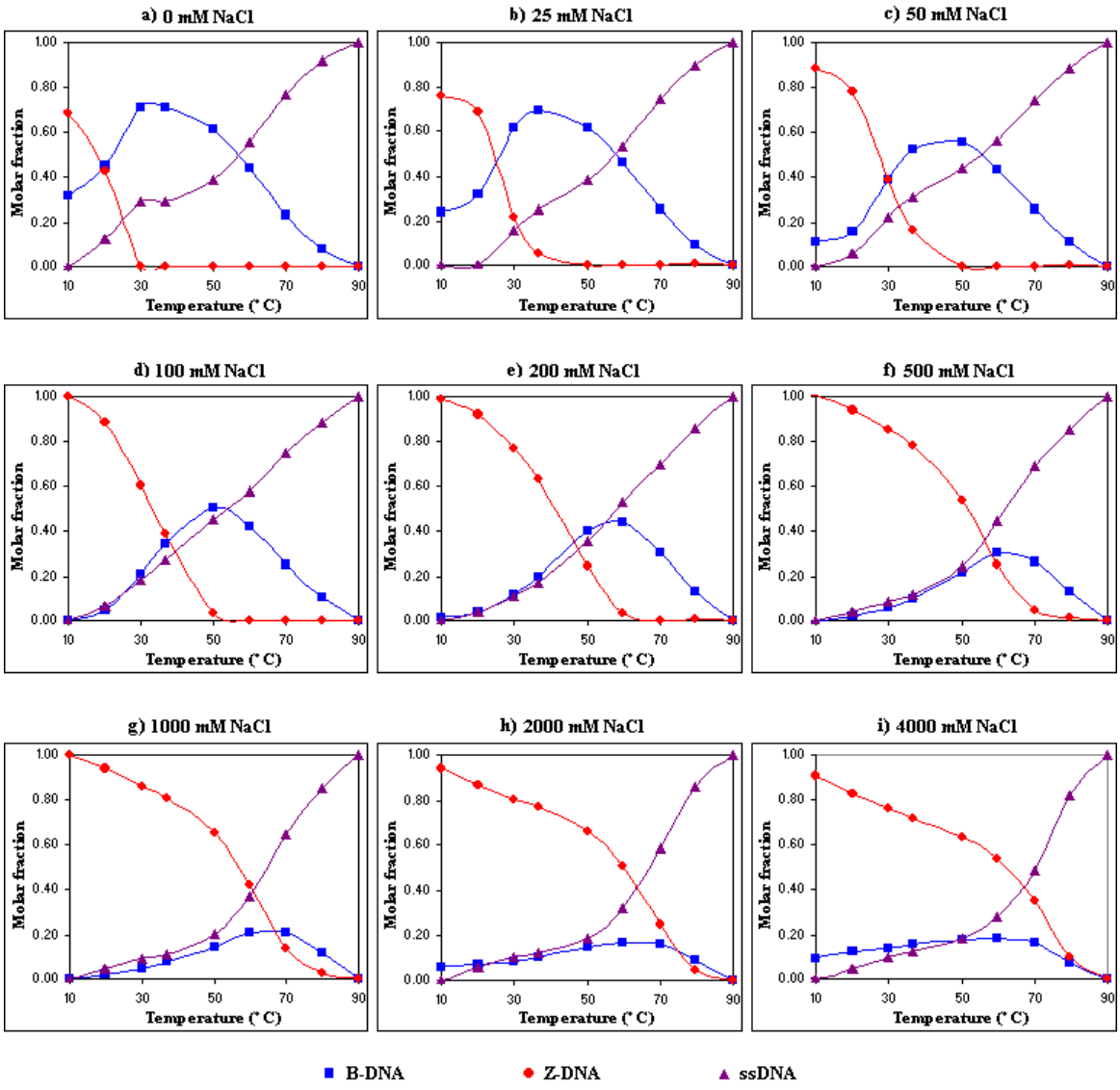


Figure 4.10 Plots of the molar fractions of B-DNA, Z-DNA, and ssDNA of CG^{8CPh} over the temperature range 10-90°C. The oligonucleotide samples were in 10 mM sodium phosphate buffer, pH 7.4 and 0 mM (a), 25 mM (b), 50 mM (c), 100 mM (d), 200 mM (e), 500 mM (f), 1000 mM (g), 2000 mM (h), and 4000 mM (i) NaCl.

Table 4.5 Molar fractions of B-DNA, Z-DNA, and ssDNA of CG^{8MMPh} with salt 0-4000 mM at temperature 10-90°C

NaCl (mM)	f_B at Temperature (°C)								
	10	20	30	37	50	60	70	80	90
0	0.96	0.93	0.94	0.95	0.92	0.80	0.56	0.25	0.00
25	0.96	0.93	0.91	0.97	0.95	0.84	0.61	0.29	0.00
50	0.78	0.86	0.91	0.92	0.90	0.81	0.60	0.30	0.00
100	0.31	0.55	0.77	0.79	0.77	0.66	0.50	0.24	0.00
200	0.24	0.37	0.65	0.74	0.75	0.67	0.49	0.25	0.00
500	0.04	0.10	0.30	0.45	0.64	0.62	0.45	0.22	0.00
1000	0.00	0.01	0.11	0.21	0.45	0.50	0.39	0.18	0.00
2000	0.00	0.03	0.08	0.12	0.25	0.32	0.30	0.14	0.00
4000	0.01	0.07	0.11	0.14	0.17	0.20	0.17	0.07	0.00

NaCl (mM)	f_Z at Temperature (°C)								
	10	20	30	37	50	60	70	80	90
0	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
25	0.04	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
50	0.22	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00
100	0.69	0.33	0.02	0.00	0.00	0.00	0.00	0.00	0.00
200	0.76	0.58	0.17	0.04	0.00	0.00	0.00	0.00	0.00
500	0.96	0.87	0.58	0.36	0.05	0.00	0.00	0.00	0.00
1000	1.00	0.96	0.81	0.67	0.31	0.07	0.00	0.00	0.00
2000	1.00	0.95	0.88	0.80	0.60	0.33	0.09	0.03	0.00
4000	0.99	0.90	0.83	0.77	0.66	0.51	0.26	0.07	0.00

NaCl (mM)	f_{ss} at Temperature (°C)								
	10	20	30	37	50	60	70	80	90
0	0.00	0.07	0.06	0.05	0.08	0.20	0.44	0.75	1.00
25	0.00	0.05	0.09	0.03	0.05	0.16	0.39	0.71	1.00
50	0.00	0.05	0.09	0.08	0.10	0.19	0.40	0.70	1.00
100	0.00	0.11	0.21	0.21	0.23	0.34	0.50	0.76	1.00
200	0.00	0.06	0.18	0.22	0.25	0.33	0.51	0.75	1.00
500	0.00	0.03	0.12	0.18	0.31	0.38	0.55	0.78	1.00
1000	0.00	0.03	0.08	0.12	0.24	0.43	0.61	0.82	1.00
2000	0.00	0.02	0.04	0.07	0.16	0.36	0.61	0.84	1.00
4000	0.00	0.03	0.05	0.09	0.17	0.29	0.57	0.86	1.00

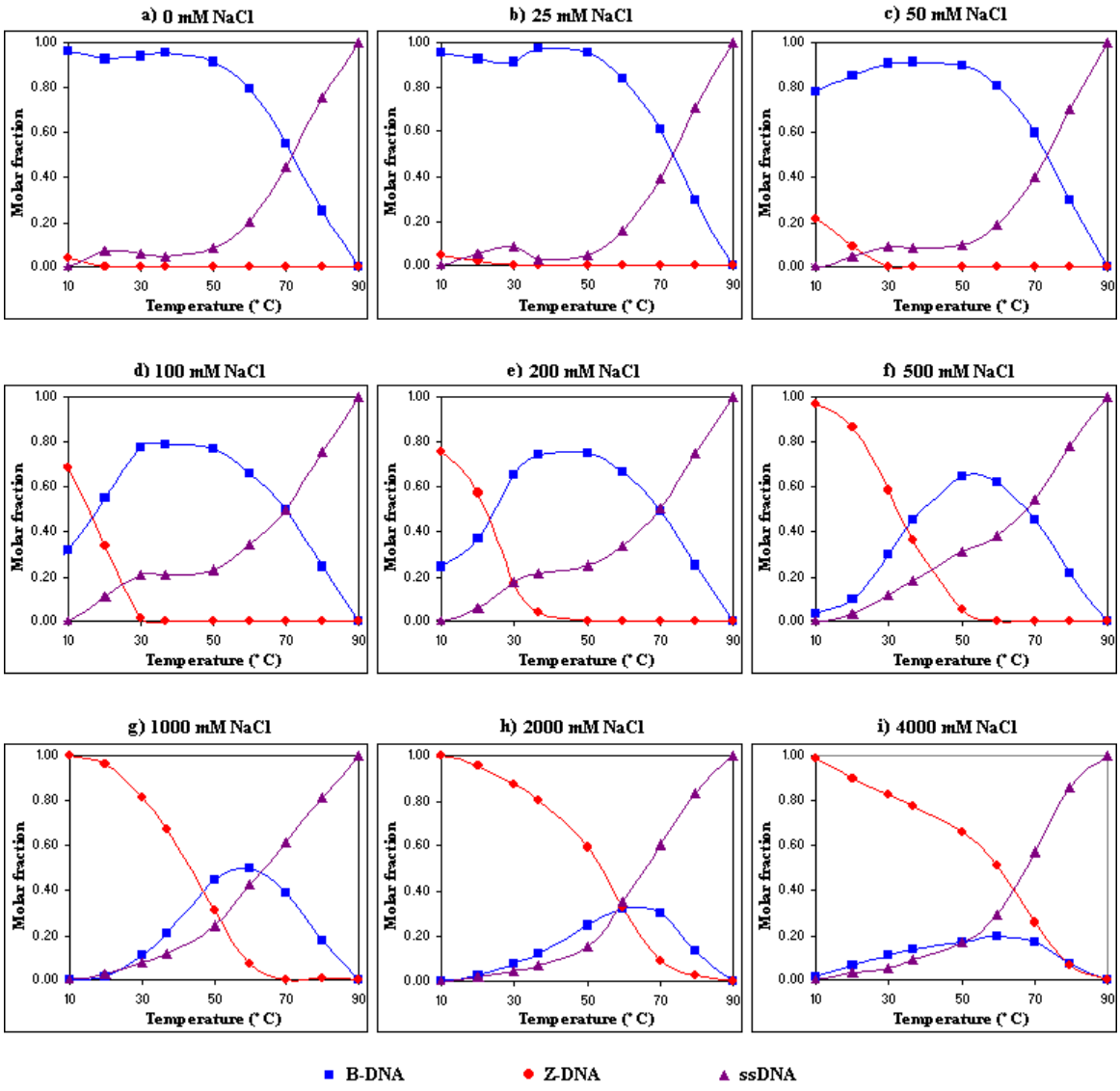


Figure 4.11 Plots of the molar fractions of B-DNA, Z-DNA, and ssDNA of CG^{8MMPh} over the temperature range 10-90°C. The oligonucleotide samples were in 10 mM sodium phosphate buffer, pH 7.4 and 0 mM (a), 25 mM (b), 50 mM (c), 100 mM (d), 200 mM (e), 500 mM (f), 1000 mM (g), 2000 mM (h), and 4000 mM (i) NaCl.

Table 4.6 Molar fractions of B-DNA, Z-DNA, and ssDNA of CG^{8HMPh} with salt 0-4000 mM at temperature 10-90°C

NaCl (mM)	f_B at Temperature (°C)								
	10	20	30	37	50	60	70	80	90
0	1.00	0.97	0.96	0.94	0.85	0.74	0.51	0.23	0.00
25	0.94	0.94	0.94	0.93	0.85	0.70	0.48	0.20	0.00
50	0.91	0.92	0.91	0.91	0.83	0.71	0.47	0.23	0.00
100	0.65	0.71	0.82	0.86	0.72	0.58	0.37	0.12	0.00
200	0.28	0.35	0.59	0.61	0.56	0.49	0.36	0.15	0.00
500	0.02	0.09	0.30	0.41	0.49	0.41	0.38	0.21	0.00
1000	0.00	0.02	0.13	0.23	0.47	0.48	0.36	0.18	0.00
2000	0.00	0.02	0.05	0.11	0.21	0.31	0.28	0.12	0.00
4000	0.04	0.10	0.15	0.18	0.23	0.24	0.21	0.08	0.00

NaCl (mM)	f_Z at Temperature (°C)								
	10	20	30	37	50	60	70	80	90
0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
25	0.06	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00
50	0.09	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00
100	0.35	0.25	0.02	0.00	0.00	0.00	0.00	0.00	0.00
200	0.72	0.55	0.12	0.02	0.00	0.00	0.00	0.00	0.00
500	0.98	0.90	0.58	0.34	0.03	0.00	0.00	0.00	0.00
1000	1.00	0.97	0.82	0.67	0.27	0.06	0.00	0.00	0.00
2000	1.00	0.95	0.88	0.80	0.60	0.33	0.09	0.02	0.00
4000	0.96	0.90	0.84	0.80	0.68	0.52	0.24	0.06	0.00

NaCl (mM)	f_{ss} at Temperature (°C)								
	10	20	30	37	50	60	70	80	90
0	0.00	0.03	0.04	0.06	0.15	0.26	0.49	0.77	1.00
25	0.00	0.03	0.06	0.07	0.15	0.30	0.52	0.80	1.00
50	0.00	0.03	0.09	0.09	0.17	0.29	0.53	0.77	1.00
100	0.00	0.04	0.15	0.14	0.28	0.42	0.63	0.88	1.00
200	0.00	0.10	0.29	0.37	0.44	0.51	0.64	0.85	1.00
500	0.00	0.01	0.12	0.26	0.48	0.59	0.62	0.79	1.00
1000	0.00	0.00	0.05	0.10	0.25	0.45	0.64	0.82	1.00
2000	0.00	0.03	0.07	0.09	0.19	0.36	0.63	0.85	1.00
4000	0.00	0.00	0.01	0.02	0.08	0.24	0.55	0.86	1.00

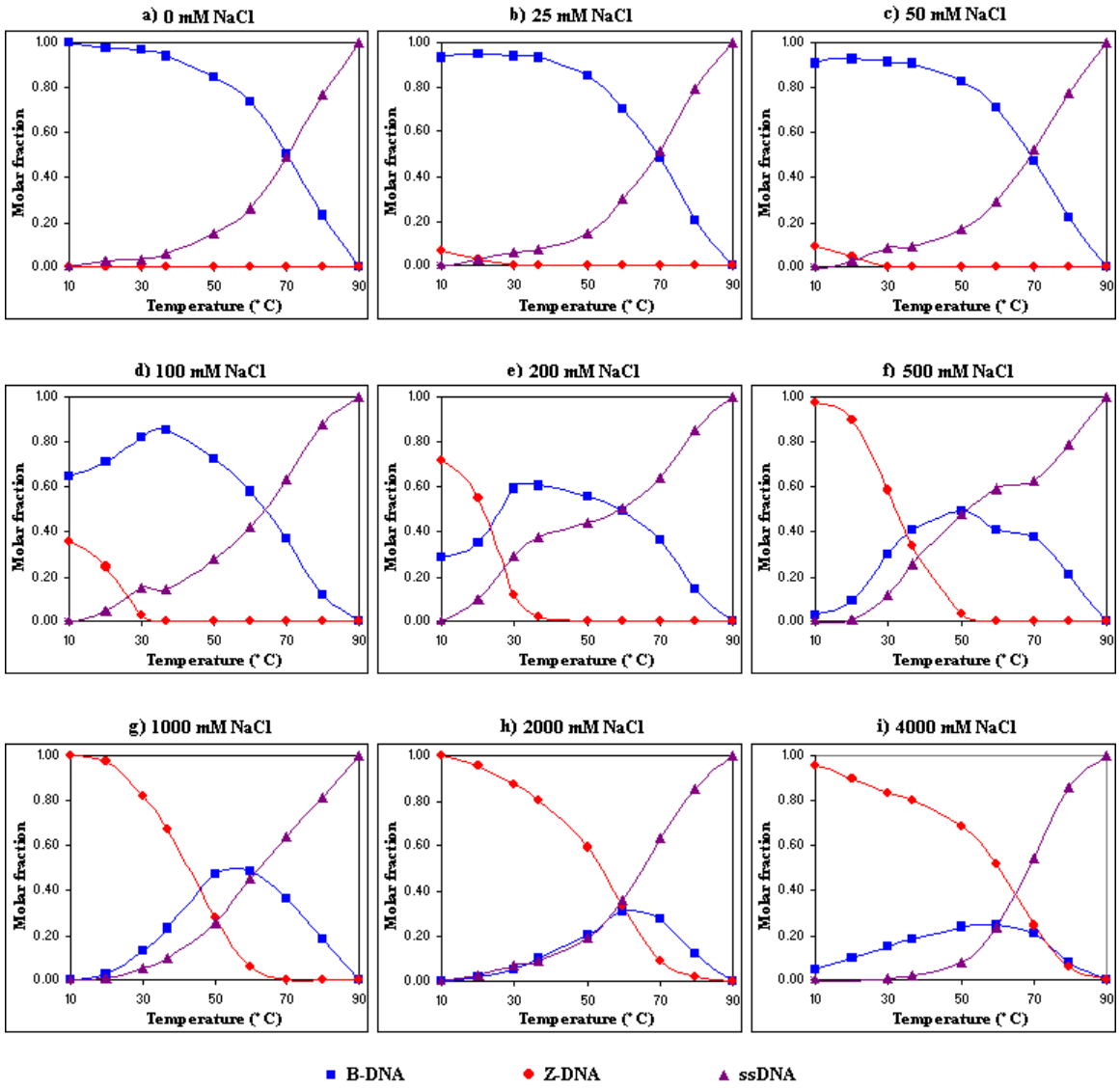


Figure 4.12 Plots of the molar fractions of B-DNA, Z-DNA, and ssDNA of CG^{8HMPh} over the temperature range 10-90°C. The oligonucleotide samples were in 10 mM sodium phosphate buffer, pH 7.4 and 0 mM (a), 25 mM (b), 50 mM (c), 100 mM (d), 200 mM (e), 500 mM (f), 1000 mM (g), 2000 mM (h), and 4000 mM (i) NaCl.

The CD spectra of the unmodified and modified CG decamers indicate that the aryl adducts formed from carcinogenic aryl hydrazines stabilize Z-DNA conformation and/or destabilize B-DNA relative to the unmodified oligonucleotide.

To quantitatively compare the B-Z transition effect among aryl adducts, the molar fraction of Z-DNA in various salt concentrations at different temperatures have been calculated from the CD data and are plotted in Figure 4.13, f_Z of CG, CG^{8Ph}, CG^{8Tol}, CG^{8CPh}, CG^{8MMPh} and CG^{8HMPh} as a function of salt concentration at 37°C. By extrapolating from the plot, to obtain 50% Z-DNA, the unmodified CG needs 3400 mM salt while CG^{8Ph}, CG^{8Tol}, CG^{8CPh}, CG^{8MMPh} and CG^{8HMPh} require 606 mM, 1000 mM, 131 mM, 669 mM, and 694 mM sodium chloride respectively. Based on these data, the B-Z transition effect of aryl adduct is ranked in the order of *p*-carboxyphenyl >> phenyl > *p*-methoxymethylphenyl > *p*-hydroxymethylphenyl >> *p*-tolyl >> unmodified

As can be seen, the mole fractions of B forms (Figure 4.14) roughly mirror the Z forms. B-DNA is predominant at lower salt concentration, while the main conformation in high salt sample is Z-DNA. Interestingly, at lower salt concentrations spikes of ssDNA formation (Figure 4.15) are seen for the modified oligonucleotides. The intensity of spikes is varied depending on the substituents on aryl adduct. This suggests that, in addition to stabilizing Z-DNA at higher salt concentration, the aryl adduct destabilizes the B form at salt concentrations less than that at which the spike occurs. This study indicates the dual effect of aryl adduct on B and Z-DNA which drives B-Z conversion.

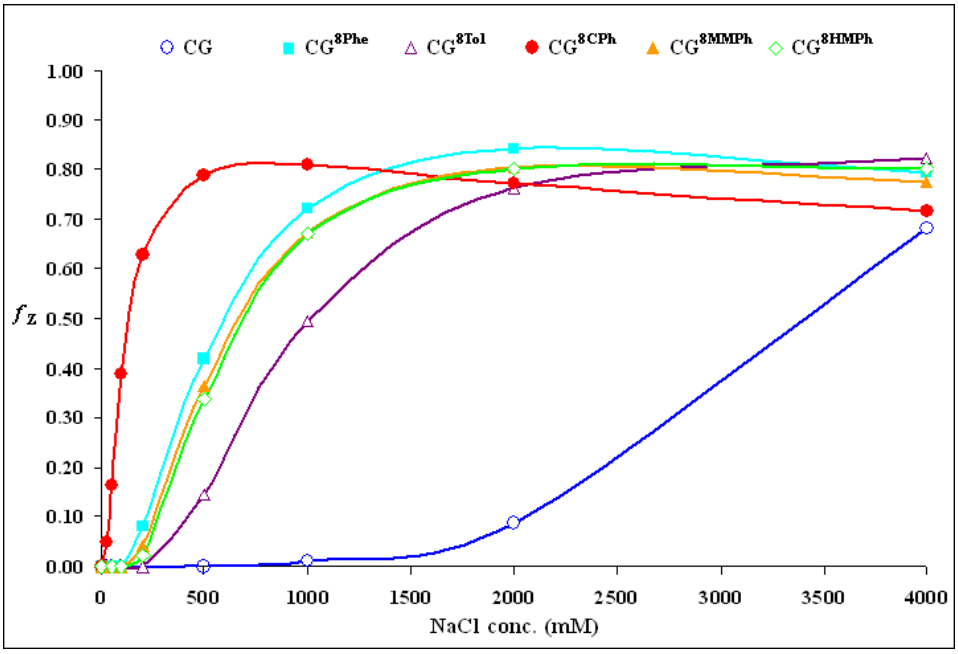


Figure 4.13 Molar fractions of Z-DNA of the unmodified and aryl modified CG decamers at 37°C as a function of NaCl concentration from 0 to 4000 mM

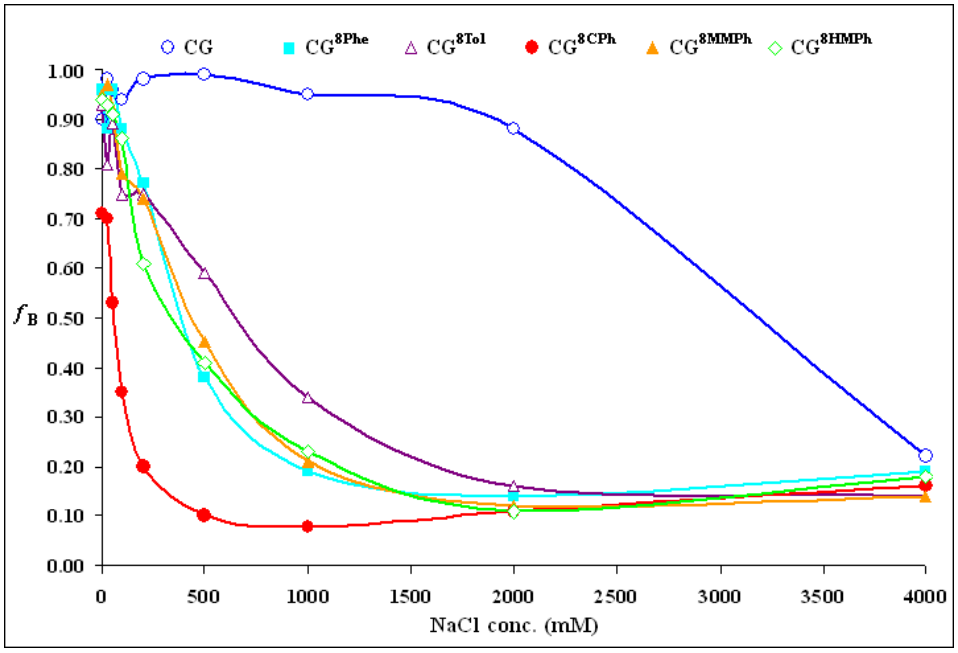


Figure 4.14 Molar fractions of B-DNA of the unmodified and aryl modified CG decamers at 37°C as a function of NaCl concentration from 0 to 4000 mM

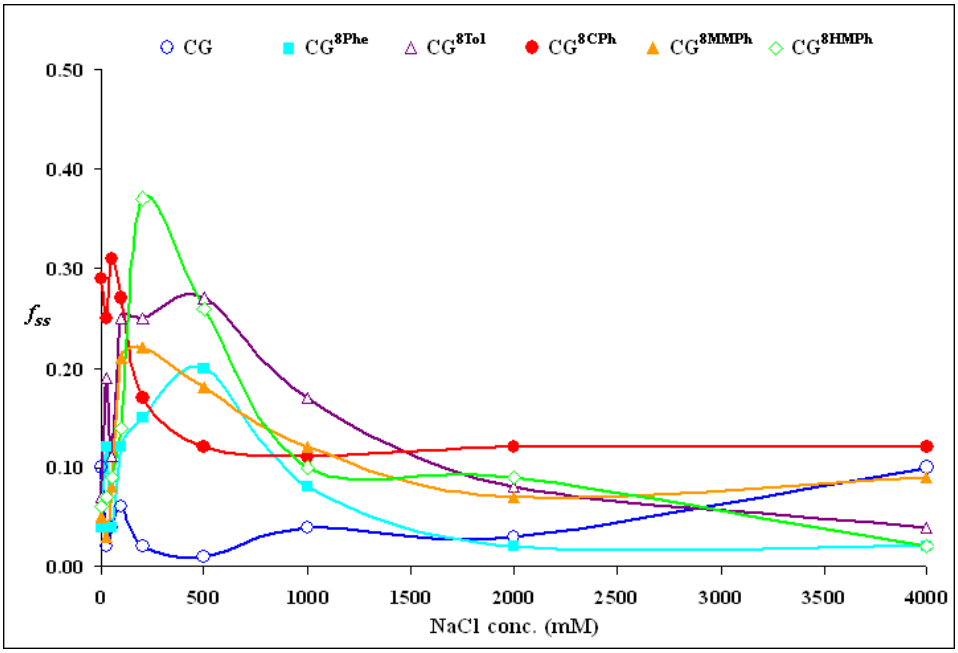


Figure 4.15 Molar fractions of ssDNA of the unmodified and aryl modified CG decamers at 37°C as a function of NaCl concentration from 0 to 4000 mM

4.2 NMR Analysis of CG Decamers Containing C8-Aryl Modified Guanine

CD is a powerful method for determining and monitoring the global conformational of DNA but can not provide information regarding the finer details of local conformation (e.g., of base-pairs). Information regarding local conformation can be obtained from NMR studies. NMR has been extensively used to study and determine the conformation(s) of nucleic acids and has relied upon two-dimensional correlation spectroscopic techniques. Overall, the technique requires the assignment of non-exchangeable proton of the synthesized oligonucleotides which is possible with the combined data from COSY and NOESY experiments as described by Scheek R.M. et al. and Orbons L.P.M. et al. In prior studies, we successfully used the assignment procedure to assign the non-exchangeable protons of CG in B form and CG^{8Ph} in Z form[22]. Though

there are potential difficulties - suitable spectra of the Z form of CG cannot not be obtained due to line-broadening cause by the high salt concentration or due to spectral complexity caused by the presence of multiple forms of DNA - the method has been shown to be sufficient to characterize the B and Z conformations of nucleic acids and we have applied it to determine the conformation of CG^{8CPh}, CG^{8HMPh}, and CG^{8MMPh}.

4.2.1 General

Deuterated phosphate buffer for NMR was prepared by drying down a known volume of phosphate buffer (10 mM, pH 7.4, 500 μL) *in vacuo* (SpeedVac) and then adding back an equal volume of deuterium oxide (D_2O, 99.996%, 500 μL). The resulting solution was then evaporated and this process repeated three times. The solid deuterated buffer salt was kept dry in cool place and was reconstituted in D_2O just prior to use.

The oligonucleotide sample (0.8 mM of CG^{8CPh}, CG^{8MMPh} or CG^{8HMPh}) was prepared in deuterated phosphate buffer (10 mM, pD 7.4, 500 μL) with or without NaCl (500 mM) and the sample solution was transferred into Shigemi NMR tube. High salt concentration was used to ensure that the Z conformation of the CG decamer was predominant in a sample. From CD data of CG^{8CPh}, CG^{8MMPh} or CG^{8HMPh} with 500 mM NaCl recorded at 30°C, the amount of Z conformation of each oligonucleotide was more than 60% which would greatly simplify the proton assignment of Z-DNA. Without NaCl, oligonucleotides are generally in B form.

All NMR spectra were measured on a Varian Unity 600 MHz spectrometer at 28°C. The ^1H NMR spectra of the oligonucleotides were obtained with solvent (D_2O) suppression. Sample concentrations were approximately 0.8 mM (10 mM, pD 7.4, 500

μL, Shigemi NMR tube was used to improve S/N) and required approximately 16 transients to obtain spectra with sufficiently high signal-to-noise. The ^{1}H-^{1}H COSY spectra were collected with 512 t_1 increments and 2048 t_2 complex points, each the sum of 16 transients. The proton nOe correlation spectra (NOESY) of the non-exchangeable protons were collected with a mixing time of 600 ms. The data were collected with 512 t_1 increments and 2048 t_2 complex points, each with the sum of 16 transients. Spectral assignments were made as described in the following sections for the B and Z DNA forms of CG^{8CPh}, CG^{8MMPh} or CG^{8HMPh}.

4.2.2 B-DNA Non-exchangeable Proton Assignment

In order to assign the non-exchangeable protons of the modified CG decamer duplexes in the B form, several types of NMR experiments are required including ^{1}H NMR, COSY, and NOESY. Basically, the 1-D proton spectrum gives a general idea of whether ss or ds DNA is present and may also provide some information regarding the DNA form (i.e., B or Z DNA). The non-exchangeable protons of DNA are typically observed in the range of 0-9 ppm. The downfield region contains base protons including H8 (δ 8.0-8.3 ppm) of guanines, H6 (δ 7.4-7.6 ppm), and H5 (δ 5.5-6.0 ppm) of cytosines are usually seperate well for CG at low salt concentration. In the present case, the aryl adduct protons (δ 6.3-7.3 ppm) are also included in this region. The sugar protons are upfield of the base protons and, for a given type of sugar proton, generally exhibit poor dispersion resulting in several clusters of unresolved peaks. Thus, to make specific assignments requires 2-D NMR techniques.

COSY data is used for the assignment of both base and sugar protons. In the case of the base protons, only the cytosine H5/H6 protons can be assigned as no other base bears protons that are adjacent to one another. Here, the COSY spectra also can be used to assign the protons of the C8-aryl group (the protons that are *ortho*- and *meta*- to the carbon attached to the modified guanine G6). In addition, the H-1′/H-2′/2″ cross peaks are fairly easy to assign by COSY, mainly because they often have a relatively broad chemical shift range. However, the assignment of the base or sugar protons to specific residues in an oligonucleotide can not be achieved using 1-D or COSY data alone and requires inclusion of NOESY spectra.

The connection between base protons and sugar protons in the same nucleoside residue can be determined from NOESY spectra due to close spatial proximities between the proton nuclei. The nOe cross peaks of Aryl/H-1′, Aryl/H-2′/2″, H8/H-1′, H8/H-2′/2″, H6/H5, H6/H-1′, and H6/H-2′/2″ in the same nucleotide subunit are usually observed in the spectra. To assign protons in and to successive bases, the connections between adjacent bases are first located. As described in the proton assignment for $d(CGC)_2$ in the B form by Orbons L.P.M. et al.[95], nOe cross peaks between H8 of G(n) and H6, H-1′, or H2′/2″ of C(n-1) were observed as well as nOes between H8, H-1′, or H2′/2″ of G(n) and H6 or H5 of C(n+1). Thus, by locating these nOe correlations they can be used as connecting points and, in turn, for assignment of protons to specific bases.

Figure 4.16 shows a diagram that schematically shows the nOe correlations that are generally used for protons assignment for oligonucleotides in the B-DNA conformation and, in particular, for a B-DNA of CG decamer. The C1 residue can often be used as the starting point because of its distinct H6 chemical shift that is the furthest

downfield as compared to the remaining cytosines (due to base-pair fraying). Then, given this assignment, the NOESY spectrum is used to identify H8 of G2 (C1-H-1′ and/or H-2′/2″ correlation with G2-H8) and, in turn, G2-H-2′/H2″ (G2-H8 correlation with G2-H-1′/2′/2″). With G2-H-1′/2′/2″ assigned, C3-H6, and in turn C3-H-1′/2′/2″ can be assigned. Continuing this process leads to assignment of all G-H8 and C-H6 protons and the H-1′/2′/2″ of the attached deoxyribose. The protons of modified G6 are also quite easy to assign due to correlations with the unique aryl protons present in the DNA adduct. Finally, the H-3′, H-4′, and H5′/5″ assignments of each base are possible by following the nOes and *J*-correlation cross peaks in NOESY and COSY spectra from the H-2′/2″ protons.

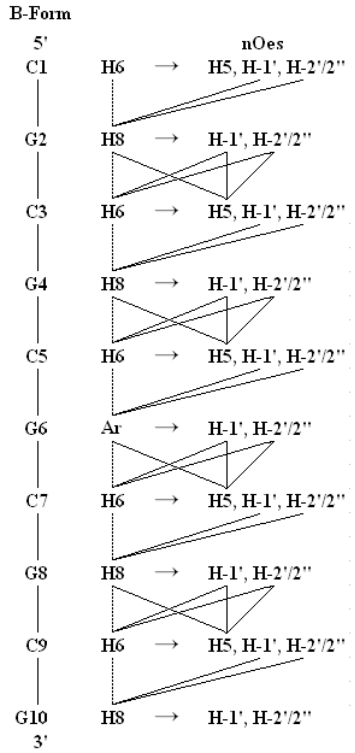


Figure 4.16 Diagram shows proton assignment strategy for B-form of CG decamers based on nOe correlations observed in NOESY spectra

To demonstrate the assignment procedure, the assignment of the non-exchangeable protons of CG^{8HMPh} will be described. The analysis begins with examination of ^1H NMR spectrum from δ5.4-8.3 ppm, shown in Figure 4.17. The four furthest downfield singlets were assigned to the H8 protons of the four unique guanines (δ8.33, 8.093, 8.089, and 8.05 ppm). Next, the cluster of neighboring upfield doublets should be the proton resonances for the five unique cytosine H6 protons (δ7.75, 7.62, 7.52, 7.50, and 6.53 ppm). The doublets at 7.09 and 6.37 were assigned to the aryl protons of G6 adduct based on their chemical shifts, and integration of about two protons which distinguishes these doublets from H6 of cytosine.

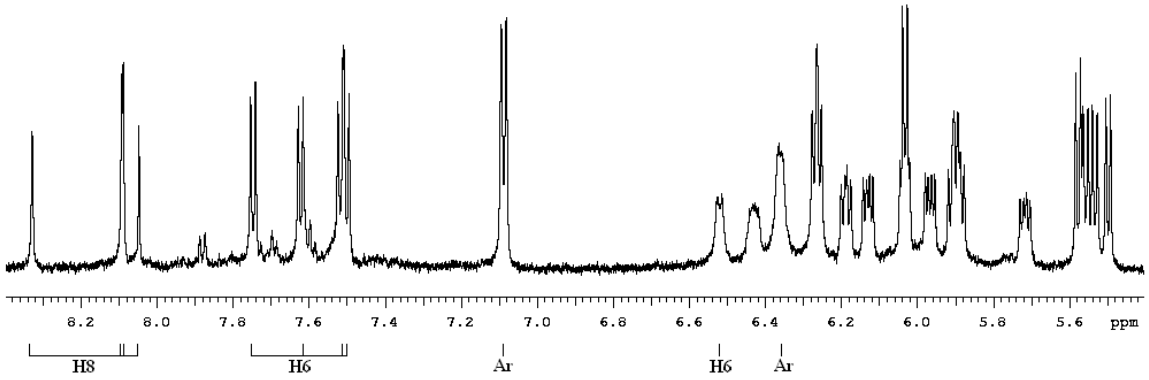


Figure 4.17 ^1H NMR spectrum of B-CG^{8HMPh} (downfield region). The resonances corresponding to H8, H6, and aryl adduct protons are noted below the frequency axis.

The COSY spectra support our assignment on the chemical shifts of H8, H6, and aryl protons. The 8.33, 8.093, 8.089, and 8.05 ppm peaks (Figure 4.18a) do not have any cross peaks with any other protons, confirming that these peaks were G-H8 protons. The peaks at 7.75, 7.62, 7.52, 7.50, and 6.53 ppm correlated to peaks at 6.03, 5.56, 5.50, 5.54, and 5.58 ppm, respectively, and arise from the *J* correlations between H6 and H5 (Figure

4.18b). As expected, the doublet at 7.09 ppm has a correltion with the peak at 6.37 ppm (Figure 4.18c) verifying that they are adjacent protons on the aryl adduct. In addition, H-1′/H-2′/2″ correlations were observed as shown in Figure 4.18d. Between 5.6 and 6.5 ppm, ten H-1′ peaks (δ5.72, 5.89, 5.91, 5.97, 6.03, 6.13, 6.19, 6.26, 6.27, and 6.43 ppm) were identified based on the criteria that these peaks had at least one correlation with peak in the region below 3 ppm which usually is the region of H-2′/2″. Most chemical shifts of H-2′/2″ were identified from H-1′/H-2′/2″ cross peaks. Next step were to assign H-1′/2′/2″ to their corresponding H8, H6/H5, or aryl protons in sequential order.

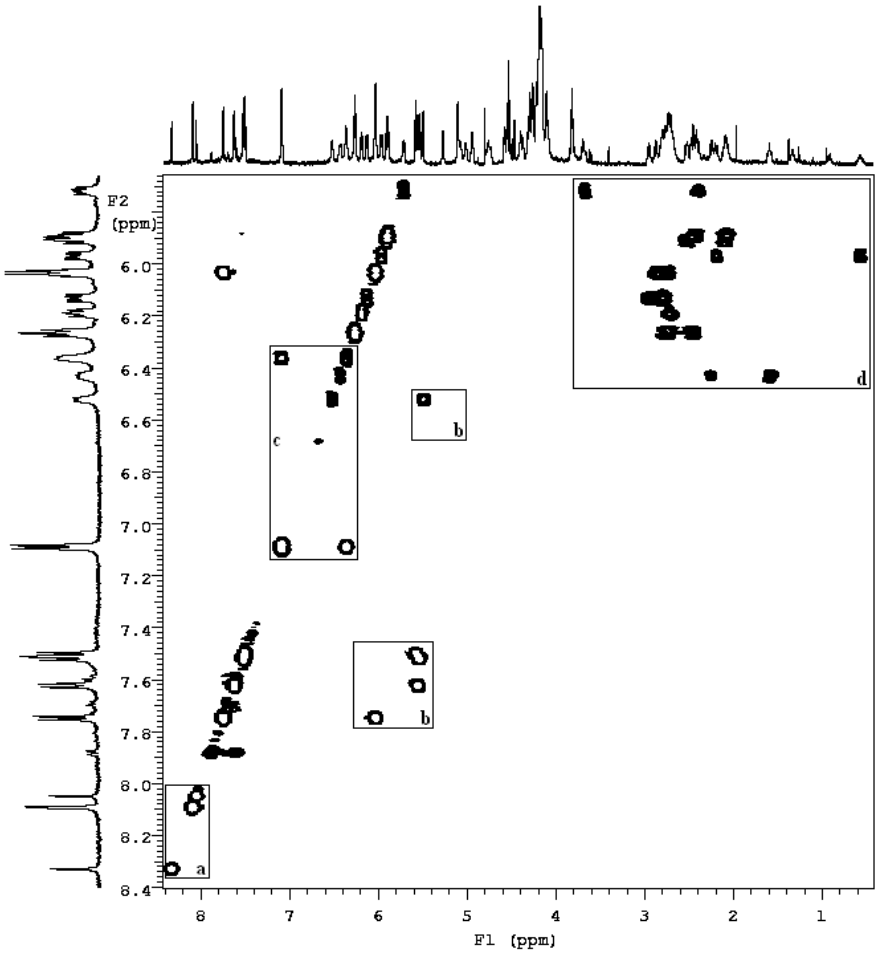


Figure 4.18 COSY of B-CG^{8HMPh} The boxed areas show a) H8, b) H6/H5, c) Ar/Ar, and d) H-1′/H-2′/2″ cross peaks.

Due to the close spatial proximity (generally from 3-5 Å based on molecular modeling) between H-1′ and H8, H6, or aryl protons in the same nucleotide residue, the nOe cross peaks of H8/H-1′, H6/H-1′, and aryl/H-1′ can be observed in NOESY spectra. Therefore, H-1′ chemical shifts can be matched up with H8 or H6 from the same bases in NOESY as shown in Figure 4.19. The assignment of H-2′/2″ from COSY can be confirmed using NOESY in similar way to H-1′ since these protons usually are close enough to have nOes with H8 or H6 in the same residue. As a result, the correlations of H8, H6, or aryl → H-1′ → H-2′/2″ were obtained including 6.53 → 5.97 → 2.20, 0.57 ppm, 7.09/6.37 → 5.72 → 3.68, 2.40 ppm, 7.50 → 5.89 → 2.43, 2.07 ppm, 7.52 → 6.43 → 2.25, 1.59 ppm, 7.62 → 6.27 → 2.69 ppm, 7.75 → 5.91 → 2.53, 2.11 ppm, 8.05 → 6.26 → 2.72, 2.46 ppm, 8.089 → 6.19 → 2.73 ppm, 8.093 → 6.13 → 2.95, 2.78 ppm, and 8.33 → 6.03 → 2.88, 2.75 ppm.

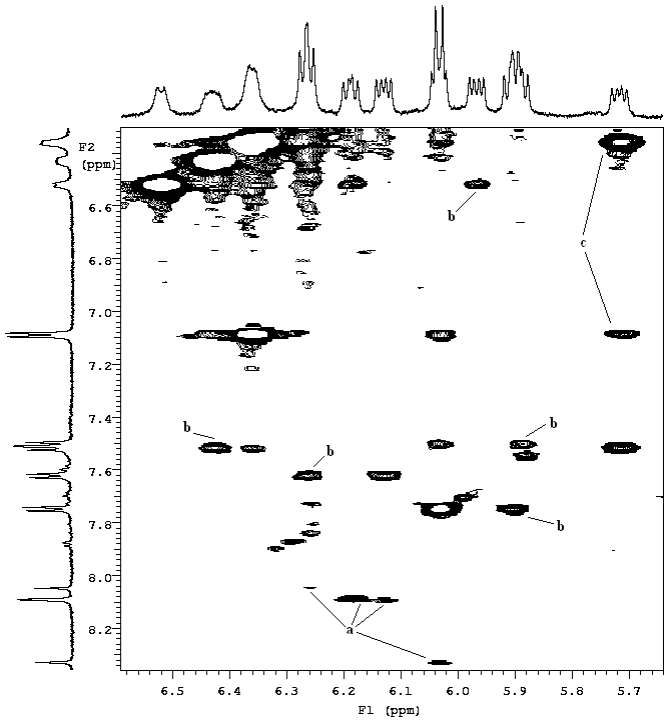


Figure 4.19 NOESY of B-CG^{8HMPh}. The areas shows a) H8/H-1′, b) H6/H-1′, c) and Ar/H-1′ correlations.

To assign the chemical shifts in sequential order, we began with the peak at 7.75 ppm which was assigned to C1-H6 due to its most downfield position among H6. The C1-H-2′/2″ (2.53, 2.11 ppm) were then used to assign G2-H8 (8.093 ppm) based on the observed nOe correlations. Likewise, G2-H-2′/2″ (2.95, 2.78 ppm) were correlated with C3-H6 (7.62 ppm). In addition, G2-H-1′ (6.13 ppm) was also correlated with C3-H6. Although C3-H-2′/2″ (2.69 ppm) was not correlated with any H8, C3 can be linked to G4 with a weak nOe from C3-H5 and G4-H8 (8.089 ppm) which allowed the further assignments of C5-H6 (6.52 ppm) through nOe with G4-H-1′/H-2′/2″ (6.19 and 2.73 ppm). The C5-H5 (5.50 ppm) has found to be correlated with aryl protons (6.37 ppm) of G6 adduct. Both G6-Ar and G6-H-1′ then correlated with C7-H6 (7.52 ppm). The nOe correlations between base(n)-H2′/2″ and base(n-1)-H8 or H6 simply allowed the assignments of G8-H8 (8.33 ppm), C9-H6 (7.50 ppm), G10-H8 (8.05 ppm) and their corresponding H-1′ and H-2′/H-2″. The uninterrupted assignment validated a previous assumption that the chemical shift of C1-H6 was 7.75 ppm. If the assignment was incorrect, an uninterrupted path would not have been observed. The NOESY in Figure 4.20 (left) shows nOe cross peaks that used for the proton assignment which are schematically shown in the diagram on the right.

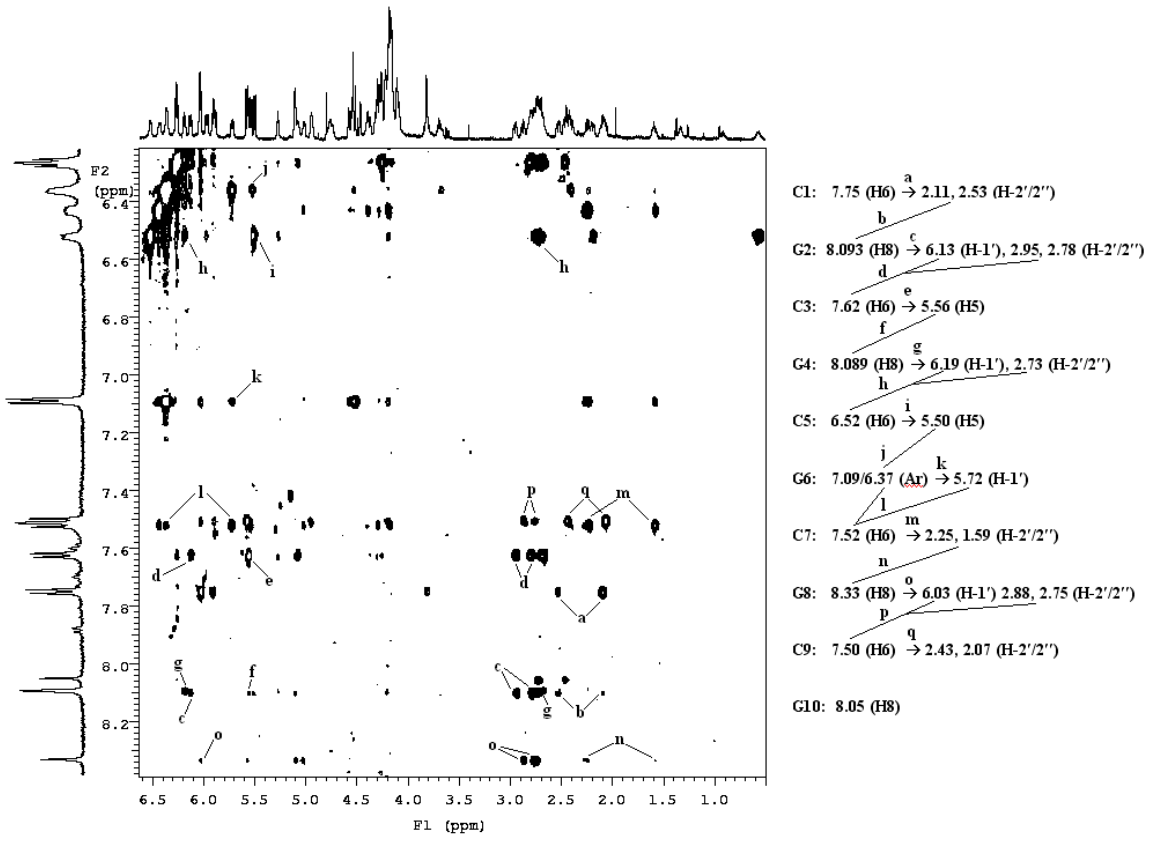


Figure 4.20 NOESY (left) of B-CG^{8HMPh} shows nOe correlations (right, a-q) that used in sequential proton assignment.

Because H-2′/2″ of each base has been identified, the H-3′ protons can be assigned based mainly on the COSY data and with a little help from the NOESY data. Due to their three bond separation, *J*-correlation between H-2′/2″ and H-3′ were observed in the COSY spectrum (Figure 4.21) and used to assign H-3′ of C1 through G10 (4.81, 5.11, 5.08, 5.27, 4.73, 4.94, 5.01, 5.10, 4.95, and 4.78 ppm, respectively).

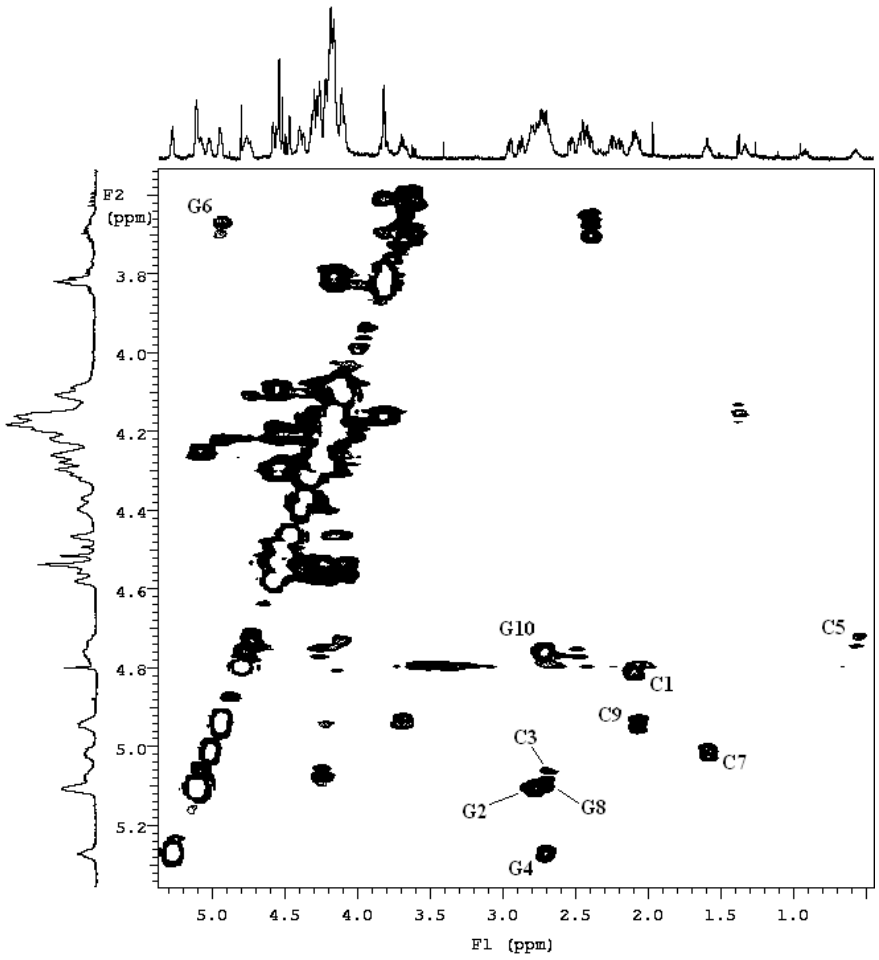


Figure 4.21 COSY of B-CG^{8HMPh}. The cross peaks between H-3′ and H-2′/2″ of C1 through G10 are marked.

Once H-3′ were assigned, H-4′ and H-5′/5″ can be obtained by simultaneous use of the NOESY and COSY spectra in the spectral regions shown in Figure 4.22. Based on the NOESY data, H-4′ and H-5′/5″ were determined from their intra-nucleotide cross peaks with H-3′, Note that nOes between H-4′ or H-5′/5″ and H-3′ of a neighboring base is insignificant or can be ruled out because of their weaker cross peak intensities (more spatial separation between nuclei) compared to the nOes of H-4′ and H-5′/5″ to H-3′ in the same base. The cross peaks found in COSY were only from H-3′/H-4′ but not from

H-3′/H-5′/5″, therefore, one can distinguish H-4′ and H5′/5″ found in NOESY spectra. By following this procedure H-4′ of C1, G2, C3, C5, G8, C9, and G10 (4.16, 4.46, 4.25, 4.11, 4.58, 4.22, and 4.26 ppm) were assigned. An alternate method was used to assign G6-H-4′ and C7-H-4′, since their H-3′/H-4′ cross peaks could not be indentified in COSY spectrum. Because H-4′ is close enough to have nOe with H8 or H6, G6-H-4′ (4.55 ppm) and C7-H-4′ (4.40 ppm) were identified based on the appearance of their H-4′/H8 and H-4′/H6 nOe cross peaks. G4-H-4′ and H-5′/5″ could not be unequivocally assigned.

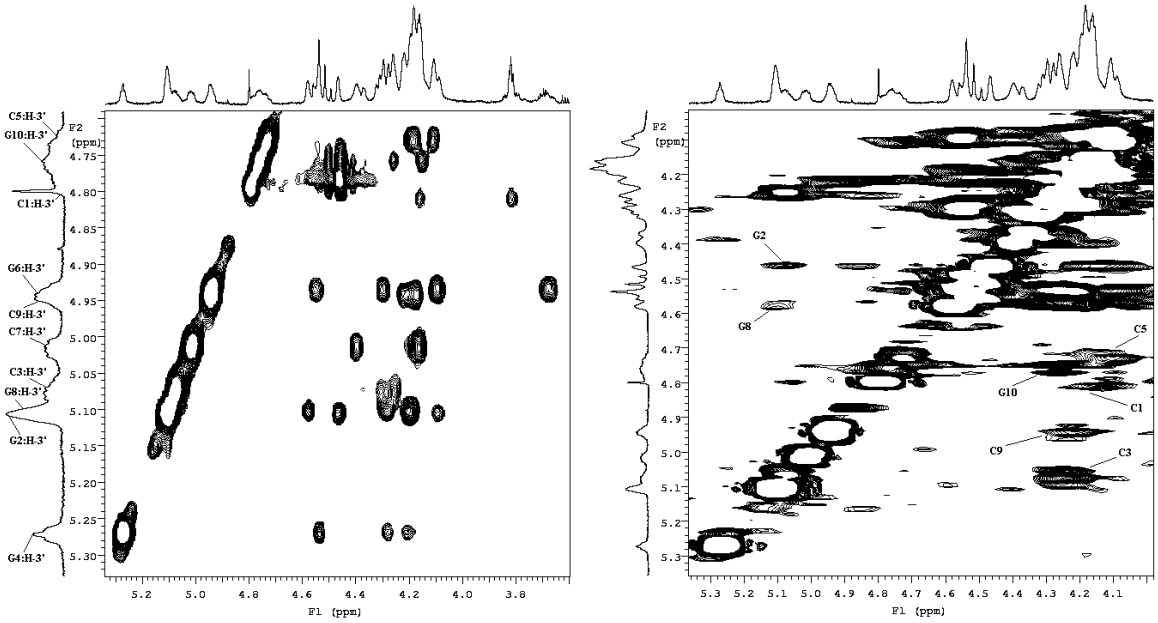


Figure 4.22 NOESY (left) of B-CG^{8HMPh} displays nOe correlation between H-3′ and H-4′/5′/5″ of all bases. On the right, COSY shows H-3′/H-4′ cross peaks of C1, G2, C3, C5, G8, C9, and G10.

By following the assignment strategy, the non-exchangeable protons of B-CG^{8HMPh} were assigned (see Table 4.7 for full chemical shift assignment). According to ^{1}H NMR, COSY, and NOESY, the patterns of spectra have verified that CG^{8HMPh} was in B conformation which is in agreement with the data from CD experiment. The non-exchangeable proton assignment of B-CG^{8MMPh} was conducted in the same way as B-

CG^{8HMPh} and is reported in Table 4.8. Unfortunately, the assignment of B-CG^{8CPh} protons was problematic due to a fairly high content of Z-CG^{8CPh} even in low salt concentration as supported by the CD data. The inability to obtain mainly B-DNA in the NMR sample significantly increases the difficulty of the NMR analysis and prevents us from successfully assigning the non-exchangeable protons for B-CG^{8CPh}. Full ^{1}H NMR, COSY, and NOESY spectra of B-CG^{8CPh}, B-CG^{8HMPh}, and B-CG^{8MMPh} can be found in Appendix C.

Table 4.7 Non-exchangeable protons assignment of B-CG^{8HMPh}

Base	Non-exchangeable ^{1}H Chemical Shifts (ppm) of CG^{8HMPh} in B Conformation[a]							
	H6/H8/Ar	H5	H1'	H2'	H2"	H3'	H4'	H5'/H5"
C1	7.75	6.03	5.91	2.53	2.11	4.81	4.16	3.81
G2	8.09	N/A	6.13	2.95	2.78	5.11	4.46	4.20, 4.09
C3	7.62	5.56	6.27	2.69	N/D	5.08	4.25	4.37, 4.30
G4	8.09	N/A	6.19	2.73	N/D	5.27	4.54	4.28, 4.20
C5	6.53	5.50	5.97	2.20	0.57	4.73	4.11	4.19
G6[b]	7.09, 6.37	N/A	5.72	3.68	2.40	4.94	4.55	4.30, 4.09
C7	7.52	5.54	6.43	2.25	1.59	5.01	4.40	4.17
G8	8.33	N/A	6.03	2.88	2.75	5.10	4.58	4.28, 4.20
C9	7.50	5.58	5.89	2.43	2.07	4.95	4.22	4.18
G10	8.05	N/A	6.26	2.72	2.46	4.78	4.26	4.15

[a]N/A – not applicable, N/D – not determined.
[b]Chemical shift of *p*-substituent on the aryl adduct; 4.58 ($-CH_2$)

Table 4.8 Non-exchangeable protons assignment of B-CG^{8MMPh}

Base	Non-exchangeable ^{1}H Chemical Shifts (ppm) of CG^{8MMPh} in B Conformation							
	H6/H8/Ar	H5	H1'	H2'	H2"	H3'	H4'	H5'/H5"
C1	7.75	6.04	5.91	2.54	2.10	4.82	4.16	3.82
G2	8.10	N/A	6.13	2.95	2.81	5.11	4.47	4.20
C3	7.62	5.56	6.28	2.69	N/D	5.08	4.26	4.38, 4.31
G4	8.09	N/A	6.19	2.72	N/D	5.27	4.54	4.28, 4.21
C5	6.52	5.49	5.96	2.19	0.58	4.73	4.11	4.19
G6[b]	7.08, 6.38	N/A	5.72	3.70	2.42	4.94	4.54	4.30, 4.10
C7	7.48	5.49	6.44	2.24	1.63	5.03	4.40	4.48, 4.17
G8	8.33	N/A	6.03	2.88	2.74	5.11	4.58	4.20
C9	7.50	5.57	5.88	2.44	2.08	4.95	4.21	4.48, 4.16
G10	8.05	N/A	6.27	2.71	2.48	4.77	4.28	4.16

[a]N/A – not applicable, N/D – not determined.
[b]Chemical shift of *p*-substituent on the aryl adduct; 4.59 ($-CH_2$) and 3.41 ($-OCH_3$)

4.2.3 Z-DNA Non-exchangeable Proton Assignment

The same basic approach as used to assign the proton resonances for the B forms of the C8-arylguanine modified oligonucleotides is used to assign resonances in the Z form. Thus, the same data set needs to be acquired, including ^{1}H NMR, COSY, and NOESY. As in the case of the B form, the 1D proton spectrum gives provides information as to whether B, Z or single-stranded DNA is present. The chemical shift ranges for the base and sugar protons are roughly the same and the peaks for H8 (δ 7.8-7.9 ppm) of guanines, aryl adduct protons (δ 7.7-8.1 ppm), H6 (δ 7.4-7.5 ppm) and H5 (δ 5.2-5.3 ppm) of cytosines can be seen. However, as can be seen, the resolution is not as good as for the B DNA spectra. This is due to the line broadening caused by higher salt concentration, required for the oligonucleotide to adopt the Z conformation. The *J* correlations of H6/H5 can be observed from COSY spectra as well as the correlation between the aryl protons that are *ortho*- and *meta*- to the guanine C8 carbon of the aryl adduct on G6. In addition, the correlation peaks between H-1′ and H-2′/2″ are also shown

in COSY. Based on the well defined regions on 1D ^{1}H NMR and 2D COSY spectra, the chemical shifts of aryl, H8, H5, H6, H-1′, H-2′, and H-2″ protons can be obtained. To proceed on the proton assignment, the connection between base protons and sugar protons in the same nucleoside residue are observed from NOESY spectra due to close spatial proximities between the proton nuclei. The nOe cross peaks of Aryl/H-1′, Aryl/H-2′/2″, H8/H-1′, H8/H-2′/2″, H8/H-3′, H8/H-4′, H8/H-5′/5″, H6/H5, H6/H-1′, and H6/H-2′/2″ in the same nucleotide subunit are generally observed in the spectra and can be used to assign some of sugar protons for each base. Strong nOes cross peaks of H8/H-1′ indicate that dG are in *syn* conformation which associate with structural feature of Z-DNA[20,97]. To sequentially assign protons, the connections between adjacent bases are required. Instead of weak nOe correlations between H8 of G(n) and both H6 of C(n-1) and C(n+1) observed in B form, NOESY of Z form shows relatively strong nOe cross peaks only between H8 of G(n) and H6 of C(n-1) which would quickly help lining up G(n) and C(n-1). To completely assign the base positions in CG decamer sequence, C(n+1) can be linked to G(n) through nOes cross peaks between H-4′ or H-5′/5″ of C(n-1) and H8 of G(n) as described in a Z-DNA protons assignment reported by Orbons L.P.M. et al[95].

As was the case for the assignment of protons for the oligonucleotides in the B form, the C1 base serves as the starting point because of its distinct H5 and H6 chemical shifts that are most downfield compares to the rest of cytosines due to C1 position in the sequence. Figure 4.23 shows a diagram described nOe correlations that generally used for Z-DNA assignment of CG decamer. The protons of modified G6 are also quite simple to assign due to its correlation with unique aryl protons from the DNA adduct. The H-3′, H-

4′, and H5′/5″ assignments of each base are possible by following nOes and *J*-correlation cross peaks in NOESY and COSY spectra.

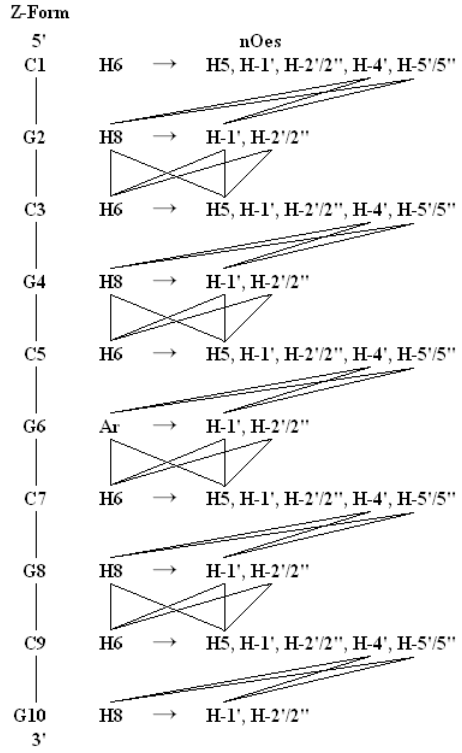


Figure 4.23 Diagram shows proton assignment strategy for Z-form of CG decamers based on nOe correlations observed in NOESY spectra

To demonstrate the assignment of the non-exchangeable protons of the modified CG decamer in Z conformation, the proton assignment of Z-CG^{8CPh} will be used as an example. The NMR spectra analysis was started on an expansion of ^{1}H NMR spectrum from δ7.0-8.3 ppm (Figure 4.24) where peaks of H8, H6, and aryl protons generally locate. The two doublets at 8.16 and 7.81 ppm were expected to be aryl protons of G6 adduct based on their chemical shifts, and integration of about two protons. The next four singlet peaks were suggested to be H8 of guanines (δ7.91, 7.88, 7.87, and 7.85 ppm),

while the cluster of peaks around 7.4-7.5 ppm are assigned to the cytosine H6 protons. Because the resolution of spectrum was fairly poor due to the high salt concentration in NMR sample, the H6 cytosine protons could not be clearly assigned at this point.

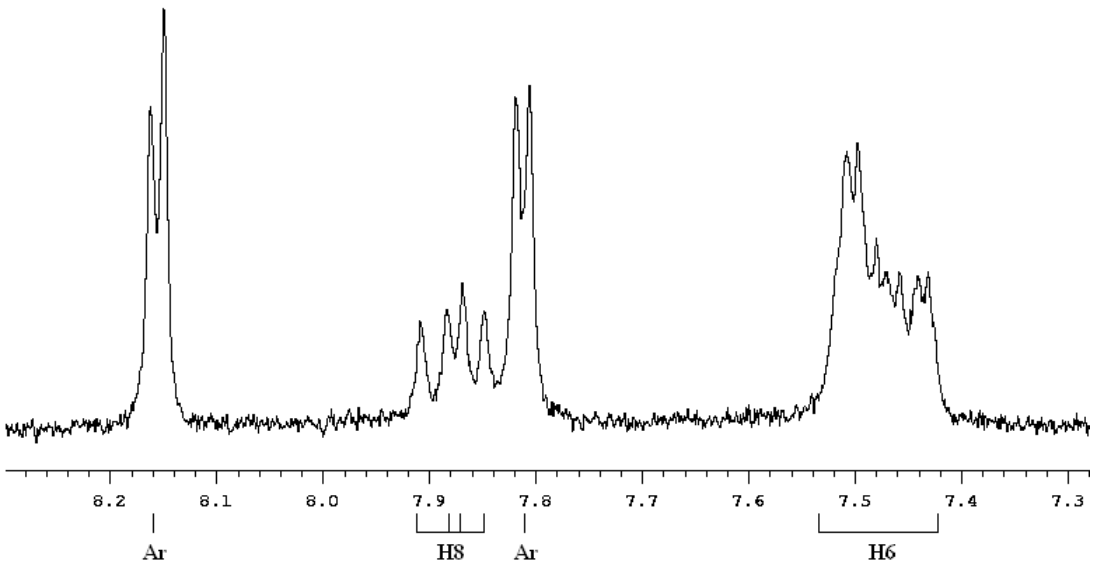


Figure 4.24 ^{1}H NMR spectrum of Z-CG^{8CPh}. The peaks corresponding to H8, H6, and aryl adduct protons are indicated below the frequency axis.

The COSY spectra (Figure 4.25a) support our suggestion on the chemical shifts of H8 (Figure 4.25a) and aryl protons (Figure 4.25c). The peaks at 7.91, 7.88, 7.87, and 7.85 ppm did not have cross peaks with any others which suggests that these peaks are the of H8 protons. As expected, the doublet at 8.16 ppm is correlated with the doublet at 7.81 ppm (Figure 4.25b) verifying that they were protons of the aryl group. Because of mesomeric effect, the downfield doublet at 8.16 ppm was assigned to the aryl protons located *ortho* to the electron withdrawing *p*-carboxy substituent, while the upfield doublet at 7.81 ppm was assigned to the protons *ortho* to the guanine C-8 carbon. Five pairs of H6/H5 correlations (δ7.51/5.29, 7.50/5.83, 7.49/5.29, 7.46/5.19, and 7.44/5.20 ppm) were

seen in the COSY (Figure 4.25b) spectrum and were used to assign H6 and H5 of cytosines. In addition, several H-1′/H-2′/2″ correlations can be observed in COSY as shown in Figure 4.25d. Between 5.7 and 6.4 ppm, six H-1′ peaks (δ5.81, 5.87, 5.94, 6.30, 6.33, and 6.34 ppm) were identified based on the criteria that these peaks had at least one correlation with peak in the region upfield of 3 ppm which usually is the region where the H-2′/2″ (Z DNA) protons are observed. Some of the chemical shifts for H-2′/2″ were identified from H-1′/H-2′/2″ cross peaks. The assignment of H-1′/2′/2″ to their corresponding H8, H6/H5, or aryl protons were conducted in the next step followed by sorting protons in CG decamer sequence.

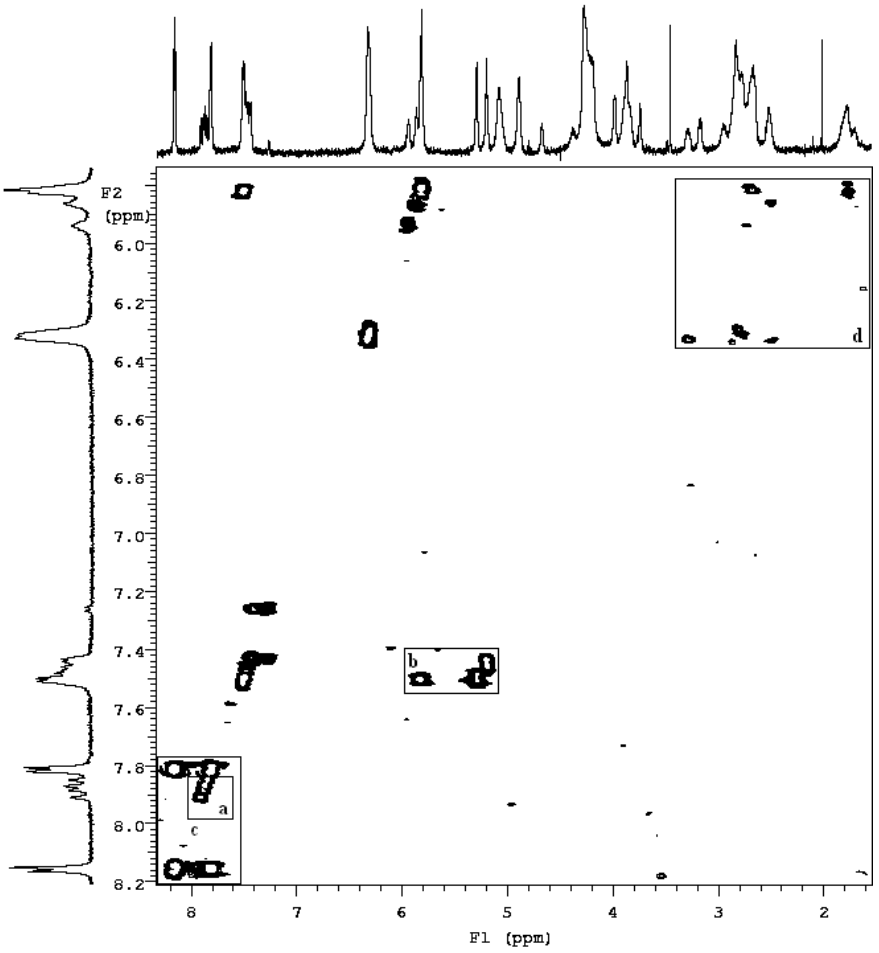


Figure 4.25 COSY of Z-CG^{8CPh}. The boxed regions show the a) H8, b) H6/H5, c) Ar/Ar, and d) H-1′/H-2′/2″ crosspeaks.

Due to the close spatial proximity between H-1′ and H8, H6, or aryl protons in the same nucleotide residue, nOes cross peaks of Ar/H-1′ (Figure 4.26a), H8/H-1′ (Figure 4.26b), and H6/H-1′ (Figure 4.26c) can be observed in NOESY spectra. It should be noted that H8/H-1′ nOes have relatively intense cross peaks because G residues were in *syn*-conformation which make H8 closer to H-1′ (2-3 Å) relative to the *anti*-conformation that all bases adopt in B form. The intense H8/H-1′ nOe was the strongest evidence indicating that the CG^{8CPh} oligonucleotide was in Z form. From COSY (Figure 4.25d) and NOESY (Figure 4.26d) spectra, H8, H6, or the aryl adduct protons were matched with their corresponding H-1′ and H2′/2″ (H8, H6, or aryl → H-1′ → H-2′/2″: 7.44 → 5.81 → 2.69, 1.77 ppm, 7.46 →5.81 → 2.69, 1.77 ppm, 7.49 → 5.81 → 2.69, 1.77 ppm, 7.50 → 5.87 → 2.51, 1.71 ppm, 7.51 → 5.94 → 2.73, 1.82 ppm, 7.85 → 6.30 → 2.83, 2.66 ppm, 7.87 → 6.33 → 3.28, 2.51 ppm, 7.88 → 6.30 → 2.81, 2.67 ppm, 7.91 → 6.34 → 2.88, 2.66 ppm, 8.16/7.81 → 6.33 → 2.94, 2.77 ppm).

Based on the computational model of Z-CG^{8CPh}, the distances between C(n-1)-H6 and G(n)-H8 are generally greater than 6 Å while the distances between C(n+1)-H6 and G(n)-H8 are below 6 Å. Therefore, only nOes of C(n+1)-H6 and G(n)-H8 are observed in the NOESY spectrum while none of C(n-1)-H6/G(n)-H8 cross peaks are observed, unlike the case of B conformation where the nOes of C(n-1)-H6/G(n)-H8 and C(n+1)-H6/G(n)-H8 may be seen. Therefore G(n) and C(n+1) can be unambiguously assigned which was a result from nOes of C(n-1)-H6/G(n)-H8. Four H8/H6 nOes cross peaks (7.85 → 7.44 ppm, 7.88 → 7.46 ppm, 7.91 → 7.51 ppm, and 8.16/7.81 (G6-Ar) → 7.49 (C7-H6) ppm, Figure 4.26e) were observed which allow the assignment of C1-H6 (7.50 ppm) and G10-H8 (7.87 ppm).

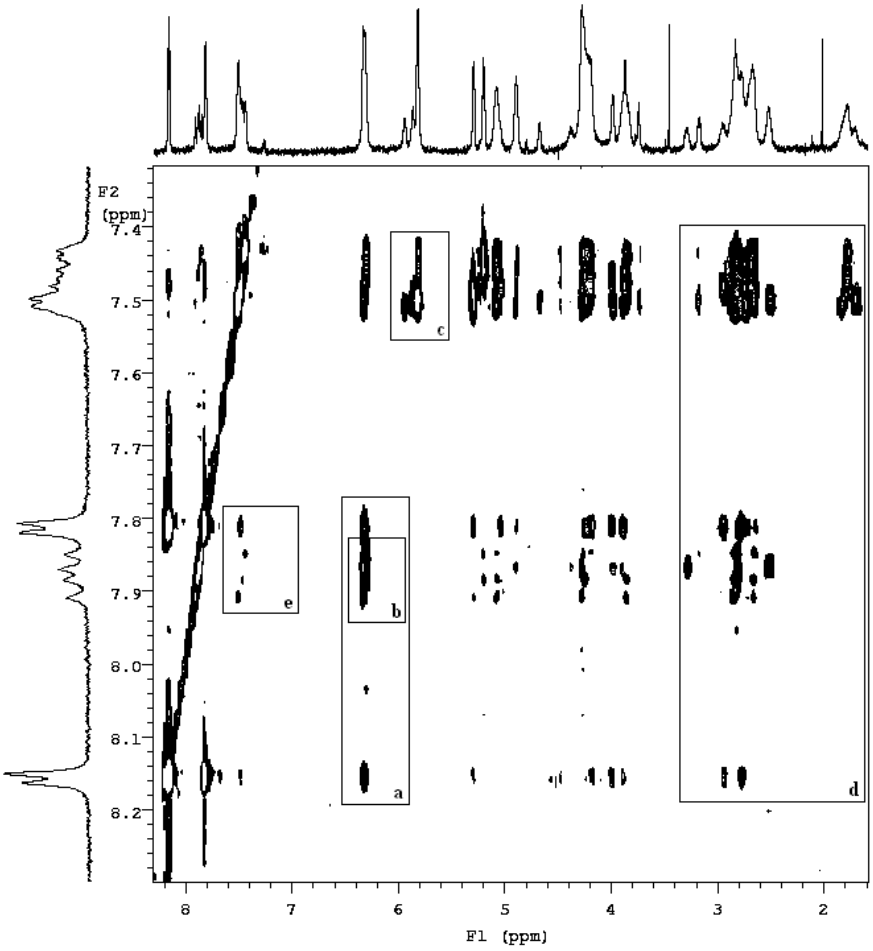


Figure 4.26 NOESY of Z-CG^{8CPh}. The boxed regions show the nOe correlations of a) Ar/H-1′, b) H8/H-1′, c) H6/H-1′, d) Ar/H-2′/2″, H8/H-2′/2″, or H6/H-2′/2″, and e) G(n):H8/C(n-1):H6.

The chemical shift of C1-H-3′ (4.67 ppm) was assigned from COSY (cross peaks of C1-H-2′/2″ and C1-H-3′) and NOESY (C1-H6 and C1-H-3′) which allow further assignment of C1-H-5′/5″ (3.17 ppm). The C1-H-5′/5″ then correlated with G2-H8 (7.85 ppm) as shown in Figure 4.27 which allowed an assignment of C3-H6 (7.44 ppm) since the nOe cross peak of G2-H8/C3-H6 has been previously found. Similarly, C3 can be connected to G4 through the cross peak between C3-H-5′/5″ (3.85 ppm) and G4-H8 (7.88 ppm) followed by an assignment of C5-H6 (7.46 ppm). The only unassigned pair of

H8/H6 has been assigned to G8-H8/C9-H6 (7.91/7.51 ppm). The H5, H-1′, and H-2′/2″ were assigned, accordingly, to match with H8, H6, and Ar assignments.

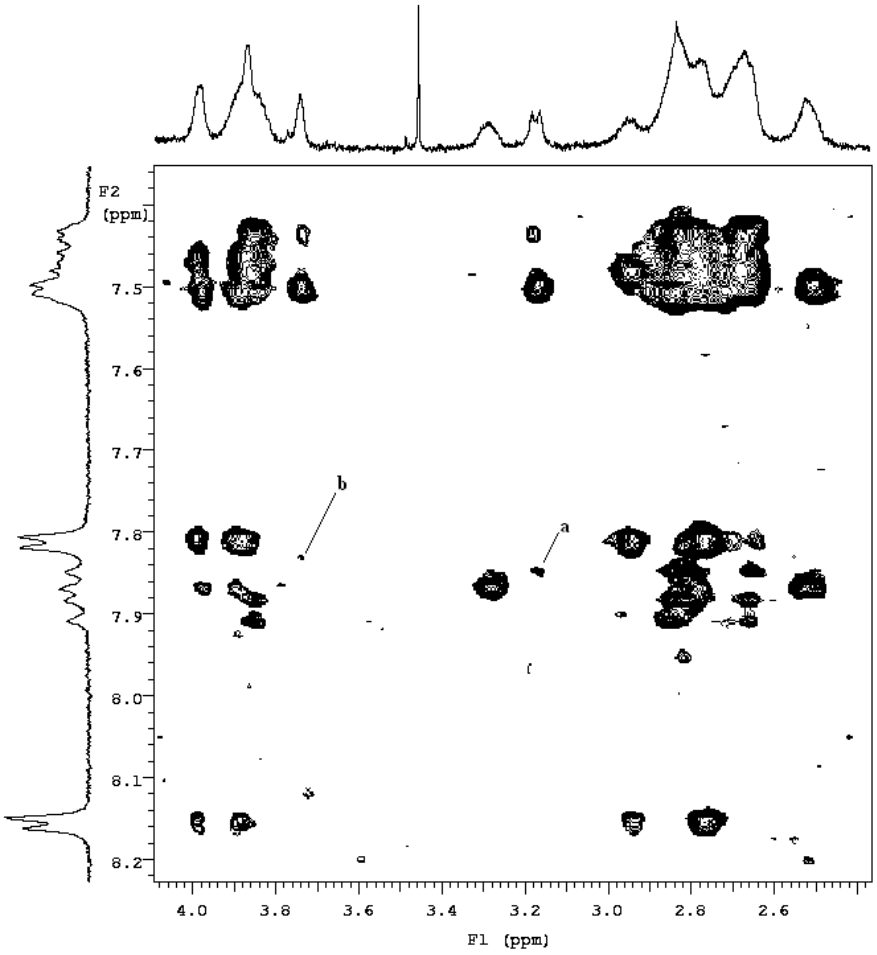


Figure 4.27 NOESY of Z-CG^{8CPh} showing the nOe correlations of a) C1:H-5′/5″/G2:H8 and b) C3:H-5′/5″/G4:H8.

Because the *J*-correlation of H-3′ and H-2′/2″ were barely observed in high salt samples, the assignment of H-3′ was made based on nOes of H8/H-3′ and H6/H-3′ given the chemical shifts of H-3′ were expected to be in range of 4.8-5.1 ppm. Unlike in B-CG^{8HMPh}, the assignment of H-4′ and H-5′/5″ was more problematic because of the

overlapping H-3′ peaks. Therefore, the chemical shifts of H-4′ and H-5′/5″ were made based on the correlation of H-4′ and H-5′/5″ with intranucleotide Ar, H8 or H6 in NOESY and are not unequivocal.

By following the assignment strategy, the non-exchangeable protons of Z-CG^{8CPh}, Z-CG^{8HMPh}, and Z-CG^{8MMPh} were assigned (see Table 4.9 to 4.11 for chemical shift assignment). According to ^{1}H NMR, COSY, and NOESY, the patterns of spectra have verified that the aryl modified CG decamers were in Z conformation which is in agreement with the data from CD experiment. Full ^{1}H NMR, COSY, and NOESY spectra of Z-CG^{8CPh}, Z-CG^{8HMPh}, and Z-CG^{8MMPh} can be found in Appendix C.

Table 4.9 Non-exchangeable protons assignment of Z-CG^{8CPh}

Base	Non-exchangeable ^{1}H Chemical Shifts (ppm) of CG^{8CPh} in Z Conformation							
	H6/H8/Ar	H5	H1'	H2'	H2"	H3'	H4'	H5'/H5"
C1	7.50	5.83	5.87	2.51	1.71	4.67	3.74	3.17
G2	7.85	N/A	6.30	2.83	N/D	5.08	4.26	4.20
C3	7.44	5.20	5.81	2.69	1.77	4.88	4.48	3.85
G4	7.88	N/A	6.30	2.81	N/D	5.08	4.24	N/D
C5	7.46	5.19	5.81	2.69	1.77	4.89	4.24	3.89
G6	8.16, 7.81	N/A	6.33	2.94	2.77	5.05	4.19	3.99
C7	7.49	5.29	5.81	2.69	1.77	4.89	4.48	3.87
G8	7.91	N/A	6.34	2.88	N/D	5.09	4.28	N/D
C9	7.51	5.29	5.94	2.73	1.82	4.90	4.48	4.29, 3.89
G10	7.87	N/A	6.33	3.28	2.51	4.90	4.38	3.99

[a]N/A – not applicable, N/D – not determined.

Table 4.10 Non-exchangeable protons assignment of Z-CG^{8HMPh}

Base	Non-exchangeable ^{1}H Chemical Shifts (ppm) of CG^{8HMPh} in Z Conformation							
	H6/H8/Ar	H5	H1'	H2'	H2"	H3'	H4'	H5'/H5"
C1	7.52	5.85	5.86	2.50	1.71	4.67	3.75	3.18, 2.84
G2	7.85	N/A	6.30	2.82	N/D	5.08	4.46	4.19, 3.74
C3	7.44	5.20	5.81	2.67	1.77	4.88	4.26	3.85, 2.83
G4	7.88	N/A	6.30	2.82	N/D	5.06	4.46	4.21, 3.97
C5	7.46	5.19	5.81	2.68	1.78	4.89	4.25	3.87, 2.82
G6[b]	7.76, 7.73	N/A	6.30	2.94	2.76	5.05	4.46	4.19, 3.98
C7	7.48	5.29	5.81	2.68	1.77	4.89	4.21	3.86
G8	7.91	N/A	6.33	2.85	N/D	5.06	4.47	3.76
C9	7.52	5.30	5.95	2.72	1.83	4.90	4.28	3.85, 2.78
G10	7.89	N/A	6.33	3.28	2.53	4.90	4.39	4.20, 3.99

[a]N/A – not applicable, N/D – not determined.
[b]Chemical shift of *p*-substituent on the aryl adduct; 4.50 ($–CH_2$)

Table 4.11 Non-exchangeable protons assignment of Z-CG^{8MMPh}

Base	Non-exchangeable ^{1}H Chemical Shifts (ppm) of CG^{8MMPh} in Z Conformation							
	H6/H8/Ar	H5	H1'	H2'	H2"	H3'	H4'	H5'/H5"
C1	7.51	5.83	5.87	2.50	1.71	4.67	3.74	3.17
G2	7.85	N/A	6.30	2.82	N/D	5.07	4.47	3.72
C3	7.44	5.20	5.81	2.68	1.77	4.88	4.26	3.85, 2.83
G4	7.89	N/A	6.30	2.81	N/D	5.08	4.47	3.99
C5	7.46	5.19	5.81	2.68	1.77	4.88	4.25	3.87, 2.82
G6[b]	7.77, 7.74	N/A	6.30	2.94	2.76	5.04	4.47	4.19, 3.98
C7	7.48	5.29	5.82	2.68	1.78	4.88	4.25	3.87
G8	7.91	N/A	6.33	2.84	N/D	5.09	4.47	3.98, 3.77
C9	7.51	5.29	5.94	2.71	1.84	4.90	4.27	3.88, 2.84
G10	7.87	N/A	6.32	3.29	2.51	4.90	4.38	4.19, 3.99

[a]N/A – not applicable, N/D – not determined.
[b]Chemical shift of *p*-substituent on the aryl adduct; 4.49 ($–CH_2$) and 3.60 ($–OCH_3$)

The results obtained from NMR analysis support the CD data that indicates the predominant of Z form of the modified CG decamers at 500 mM NaCl, where the unmodified CG could not well adopt the left handed conformation. An increase in intensity of intranucleotide H8/H-1′ nOe cross peaks when high salt concentration was introduced to the oligonucleotide samples suggests the change in conformation of dG in CG decamer sequence from *anti* to *syn* conformation which suggests the conversion of B to Z form.

The B-Z transition promoted by C8-arylguanine adducts have been confirmed by NMR and CD analysis. The *p*-substituents on the aryl adducts have significant impact on their effect on B- and Z-DNA stability, as the result has shown that different *p*-substituents give diverse degrees of Z-DNA formation. What could be gleaned from the current dara set is that B-Z conversion is dependent upon more than a local steric effect of C8-aryl adduct that would destabilize B-DNA. Additional forces that override an unfavorable electronic interaction of Z-DNA may be necessary. The interplay of several factors, including hydrophobic, electronic, steric interaction, and H-bonding that helps stabilize Z-DNA and/or destabilize B-DNA are relevant in B-Z conversion process. Molecular modeling and free energy calculation of the B and Z-forms of CG decamers were conducted to help understanding the effect of aryl adduct on DNA conformation.

CHAPTER 5
MOLECULAR MODELING OF THE B AND Z DNA FORMS OF CG DECAMERS CONTAINING C8-ARYL MODIFIED GUANINE

Molecular modeling has become a significant tool to model and study biological macromolecules. Modeling studies can provide information regarding the molecular conformation of macromolecules, such as oligonucleotides, through the application of molecular mechanical (MM) and molecular dynamical (MD) methods. More recently, molecular modeling methods have been utilized to predict the stability of macromolecules by the use of free energy perturbation (FEP) calculations. Here, we have applied both molecular modeling and FEP methods to the C8-aryl modified guanine containing oligonucleotides that were studied by CD and NMR. The aim of these studies is two-fold. First, molecular models obtained by MM/MD methods are useful for interpreting the experimental data. This is especially true for NMR NOESY data where correlations (appearance and intensity) are distance dependent. Second, computational methods allow deconvolution of the various contributors to the overall macromolecule conformation and, therefore, determination of the important factors that drive it.

Analysis of the results of MM/MD calculations of the modified and unmodified oligonucleotides studied by CD and NMR, in both the B and Z conformations is necessary because base modifications often affect and distort a regular structure of DNA. Topology, helical symmetry, and base paring of DNA all can be altered and to various degrees as a result of the modification. Therefore, the changes in helicoidal and base pairing parameters are typically used to elucidate and to explain how the C8-arylguanine adduct affects the conformation of the CG decamer locally and globally. Figure 5.1

shows diagrams of these helicoidal parameters[100] including inter-base pair, intra-base pair, and base pair-axis parameters that are considered as a part of this analysis.

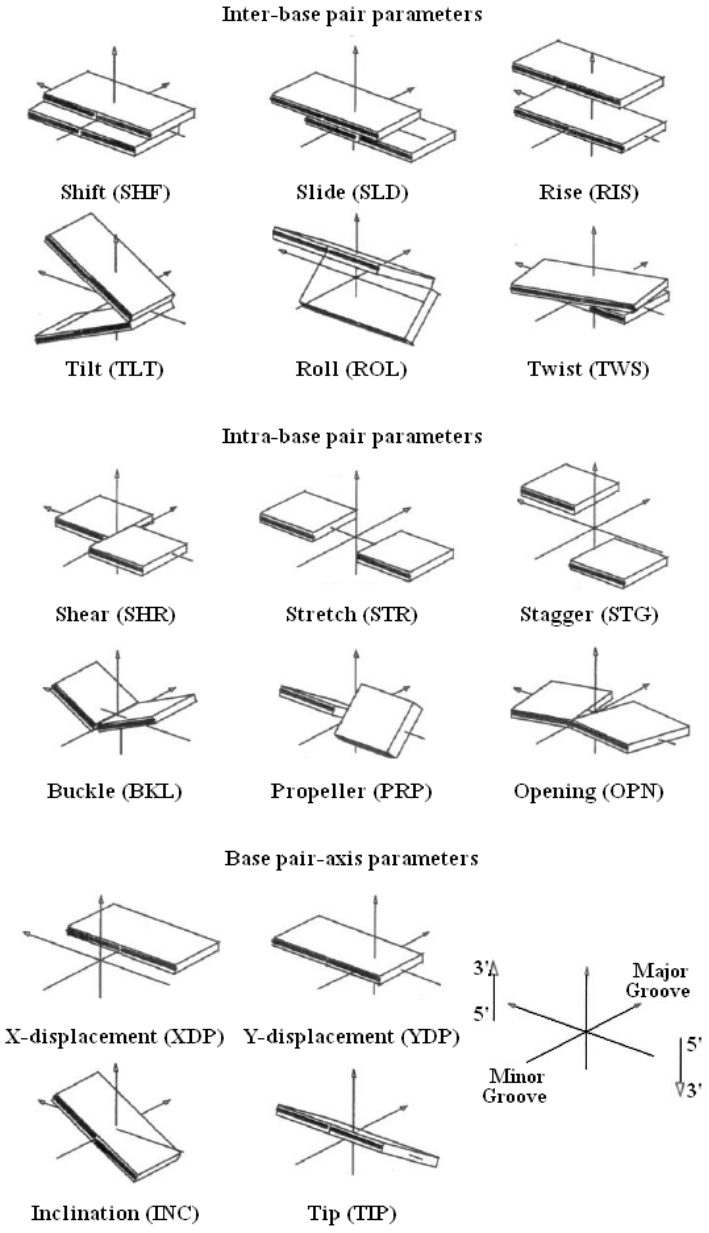


Figure 5.1 Diagrams showing the helicoidal parameters that describe the movement of base pairs and local and global conformation.

Free energy calculations of the B and Z DNA forms of the oligonucleotides result in a prediction of which form, B or Z DNA, is more favorable for a given modification. In addition to this prediction, the free energy calculations, extracted from the MD trajectories, provide a breakdown of the contributions from van der Waals interactions, electrostatics, etc., and may be used to estimate how the aryl adduct affects each component of the free energy. In turn, this may help elucidate how the aryl adducts facilitate B-Z transition, by destabilizing the B form, and/or stabilizing the Z form. It should be noted, however, that the methods used here (MM_PBSA and sietraj) are designed for protein-substrate while they binding and have been used with B-DNAs with limited success, they have not been used with Z DNAs.

The computational studies described below are a continuation of work on the 10-mers of unmodified CG, CG^{8Ph}, CG^{8Tol}, and CG^{8HMPh} conducted by Heavner[99,101] and some of these results are reproduced here for comparison purposes. The new studies refer to the CG^{8CPh} and CG^{8MMPh} derivatives. We also note that in Heavner's work, the CD and NMR work was either incomplete or not conducted for CG^{8Tol} and CG^{8HMPh} and this body of work completes Heavner's research.

5.1 General

The computational studies on CG^{8CPh} and CG^{8MMPh} generally following the protocol established by Heavner and applied to the CG, CG^{8Ph}, CG^{8Tol}, and CG^{8HMPh} decamers. The basic steps were as follows. First, parameters not provided in the parm99.dat force field that is provided with the Amber suite of programs are developed. These parameters were generated using Gaussian (g03), to calculate the natural bond

lengths and stretching constants, bond angles and bending constants, the preferential torsion angles and the atom charges. The resulting parameters were processed by antechamber to create a prep file that modifies the Gaussian output by modifying the atom names and charges to be consistent with the parm99.dat force field as well as modifications needed by the Amber (version 8[102], if not noted otherwise) modules xLEaP and sander. Subsquently, this file is examined by parmchk to screen for any remaining parameters that may be needed. The modified oligonucleotide is then built, suitable input files written for the MM calculations. MM was then run using the sander module of Amber to initially relax the oligonucleotide, solvent molecules that were added, or the entire system. MD simulations, also using sander, were then run and the resulting trajectories analyzed using utility programs contained within the Amber suite (e.g., ptraj, nmode, MM_PBSA) or other programs including Dials & Windows[103], CURVES[104] (version 5.2) , X3DNA[105], Moil-View[106] and sietraj[107,108].

5.2 Force Field Development

The structure of the C8-aryl modified dG (either CG^{8CPh} or CG^{8MMPh}) was built in Gaussview[109], then the electrostatic potentials associated with atoms in the structure were calculated using Gaussian (g03). Antechamber was then applied to parameterize C8-aryl modified dG to provide the prepi file and a pdb file. Atom types were consistent with the parm99.dat force field. The frcmod file was created by parmchk, using the prepi file as input. The pdb, prepi and frcmod files were loaded into xLEaP and a library file of the modified base generated and saved for later use. Note that in case of the C8-

carboxyphenyl dG derivative, the carboxylic proton was removed and thus was prepared as the carboxylate since this is the form it is present as under physiological conditions.

5.3 Initial Oligonucleotide Structure Building

The starting point for the B DNA forms of the two structures built here was the structure of B-CG^{8HMPh} from our previous study which was used as a template for the B-DNA form of CG^{8CPh}. The pdb file of B-CG^{8HMPh} was loaded into xLEaP using the library file for the C8-hydroxymethylphenyl dG. The aryl residue was edited, replacing the *p*-hydroxymethyl group with the *p*-carboxylate (*p*-COO$^-$) substituent and the structure was then saved as a pdb of B-CG^{8CPh}. The text editor was then used to correct the atom types, atom numbers, and assignment base of the modified dG residues (G6 and G16) in a pdb of B-CG^{8CPh} in order to comply with frcmod and lib files of C8-carboxyphenyl dG that has written previously, renaming them as CPG. The resulting pdb of B-CG^{8CPh} was an initial structure of B-CG^{8CPh}, in the deprotonated form (since it is assumed that the carboxylic group is ionized under physiological conditions) and was used as the starting structure for the MM and MD calculations. The initial structure of Z-CG^{8CPh}, B-CG^{8MMPh}, and Z-CG^{8MMPh} were constructed in the same way as described for the B-DNA form of CG^{8CPh}.

5.4 MM and MD Simulations

MM and MD were performed using the sander module of Amber and the parm99.dat version of the Cornell force field. The pdb structure was loaded into xLEaP and a water box (TIP3) was added. Charges were then neutralized by the addition of 18

(19 for CG^{8CPh}) sodium counter ions. A 10 Å buffer zone between the box wall and solute, containing solvent, was used. The pdb, topology (top), and coordinate (crd) files of CG decamers including the solvent box and sodium ions were then saved and used in the simulations.

The simulation protocol started with a MM to minimize the energy of the whole system with the DNA fixed with 500 kcal/mol Å positional restraints on all solute atoms for 1000 steps of steepest descent and then 1500 steps of conjugated gradient minimization. Then, the entire system was energy minimized (i.e. positional restraints on DNA removed) with 500 steps steepest descent and 2000 steps conjugated gradient minimization. The next step, was a 25 ps MD with the DNA fixed, was performed to equilibrate water and counter ions followed by the isothermal-isobaric ensemble (300 K°, 1 atm) 25 ps MD with the SHAKE algorithm on and a 9 Å cutoff. An additional 3 ps MD with DNA fixed was performed in order to relax water and ions around the solute. Another 600 steps MM was conducted to equilibrate solvent box with 5 kcal/mol Å positional restraints on all DNA atoms. Subsequently, the production MD with constant temperature and pressure (300 K°, 1 atm) was performed over 2 ns (4 ns for CG^{8MMPh}).

5.5 Most Representative Structure

The most representative structure is the structure that has the lowest average root mean square derivation (RMSD) to all other member of the MD trajectory. This structure is used in this work as the input into the CURVES and X3DNA routines that determine helicoidal parameters. MOIL-view was used to select the most representative structure from each of the four production trajectories. Due to the size limitations imposed by

MOIL-view, trajectories were divided into 250 ps frame pieces. Each piece of trajectory was analyzed using the cluster analysis feature in MOIL-view (2 Å cut-off distance) to select the crd set of the most representative structure for each of the 250 ps frame pieces. All of these files were then merged into a new trajectory using ptraj and the analysis repeated on this new trajectory (typically comprised of 8-16 coordinate sets). The most representative structure of the production run was generated by cluster analysis of this newly written trajectory. The pdb of most representative structure was made using ambpdb (Amber suite program) using the appropriate topology file and the coordinate set that was determined by MOIL-view to be the most representative.

In addition to using this file for determination of the helicoidal parameters, it also provides a representative model that can assist with the NMR assignment of non-exchangeable proton. The average distances between non-exchangeable protons obtained from the structure were most useful for NMR assignment in NOESY spectra. The most representative structures of all studied oligonucleotides in B forms are shown in Figure 5.2 and Z forms are shown in Figure 5.3. By inspection of the most representative structures it is obvious that the B forms were distorted by the aryl adducts. The rise and incline of base pairs in the middle of sequence, where the modified base presented, were most notable. As far as we can see by mere visual inspection of the Z forms, the aryl adducts have little effect on structure with the exception of CG^{8CPh} and CG^{8MMPh} but a more quantitative evaluation is necessary to ascertain the extent of the effect.

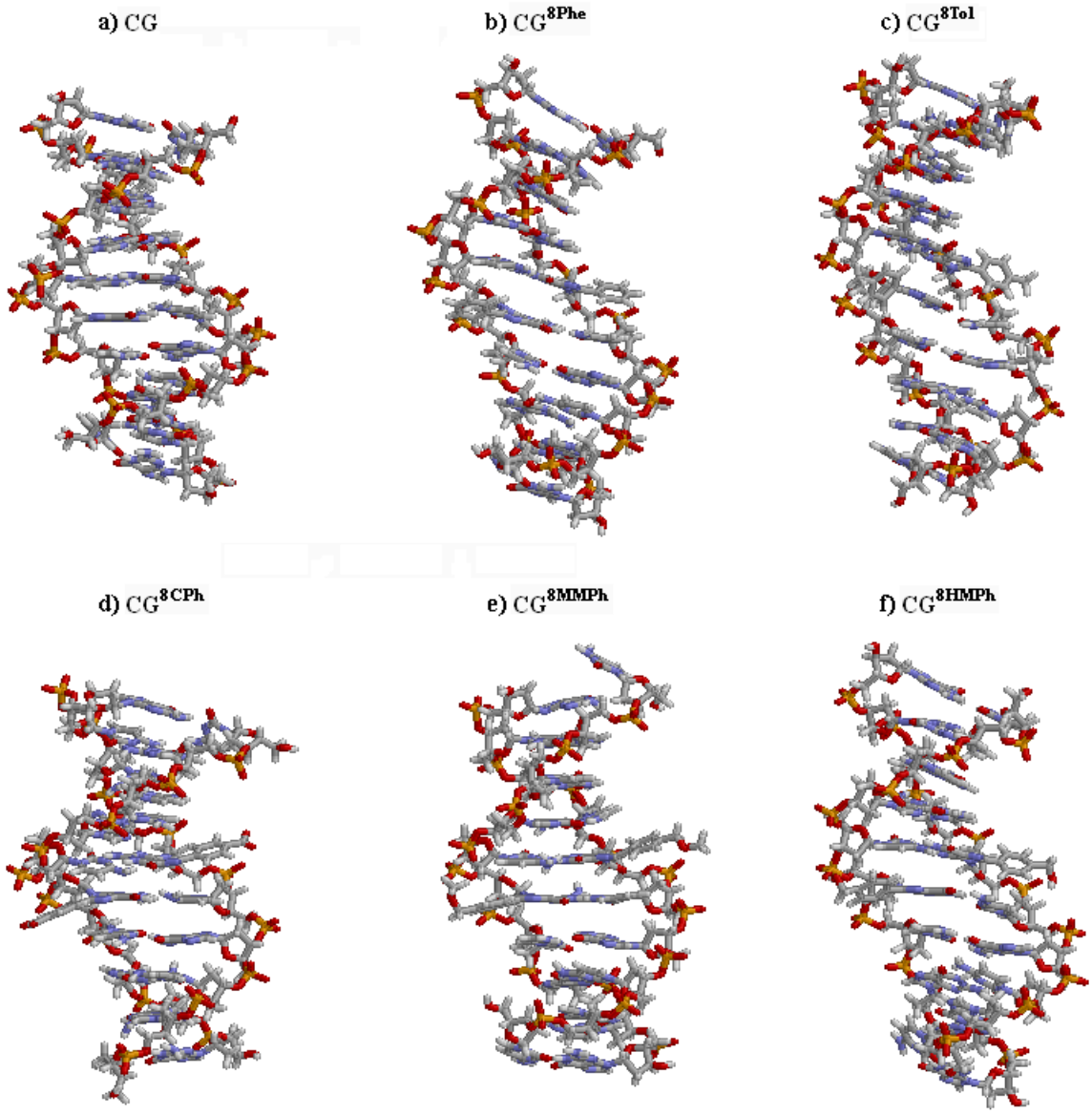


Figure 5.2 The most representative structures of a) CG, b) CG^{8Ph}, c) CG^{8Tol}, d) CG^{8CPh}, e) CG^{8MMPh}, and f) CG^{8HMPh} in the B-DNA conformation.

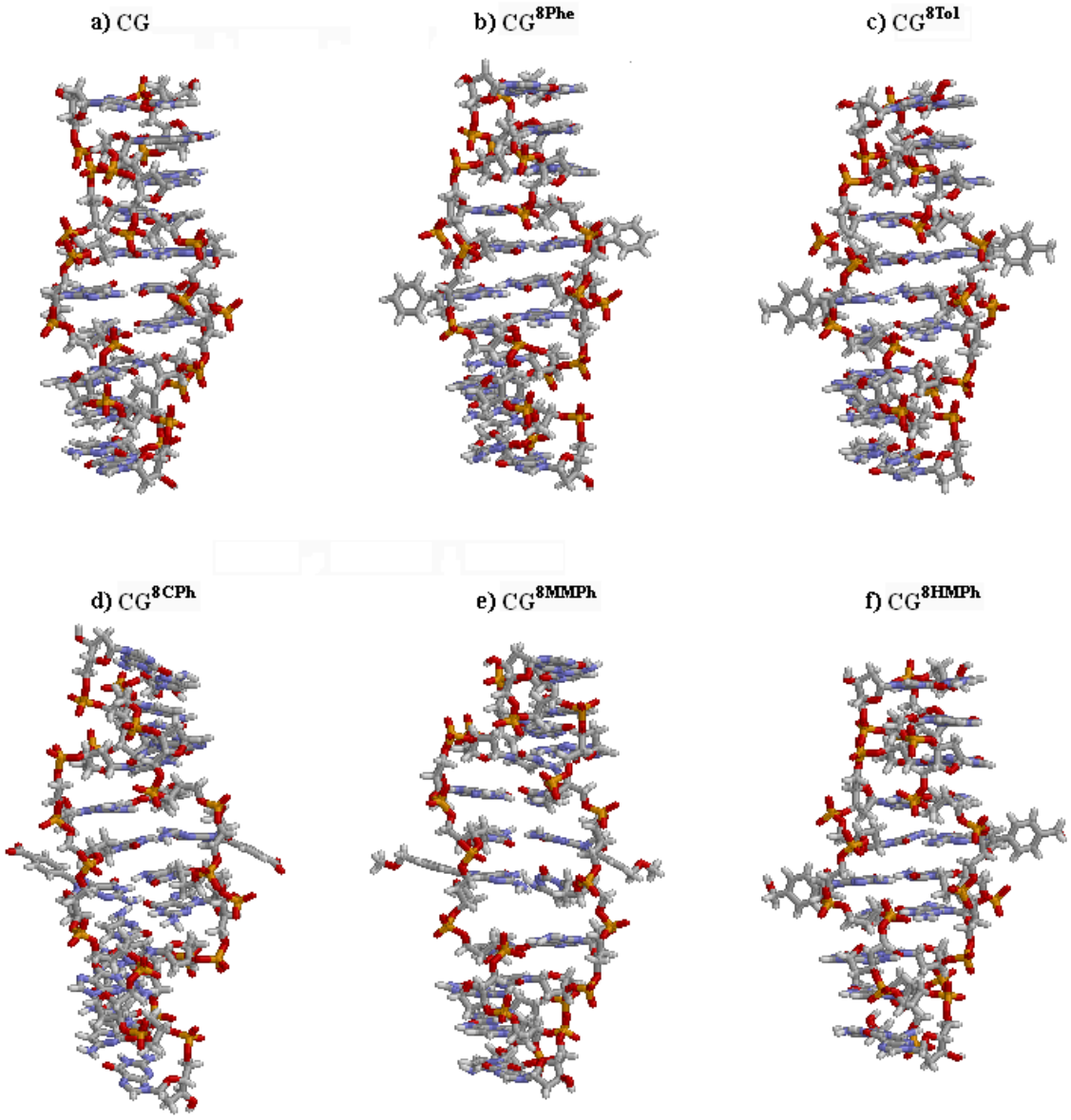


Figure 5.3 The most representative structures of a) CG, b) CG^{8Ph}, c) CG^{8Tol}, d) CG^{8CPh}, e) CG^{8MMPh}, and f) CG^{8HMPh} in the Z-DNA conformation.

5.6 Structural Analysis

5.6.1 B-DNA Oligonucleotides

Dials & Windows was used to examine the trajectories over the course of production run to determine the helicoidal parameters of the oligonucleotides that vary the most over time and with respect to the unmodified derivative. This analysis gives an excellent presentation of the variation in these parameters as the system evolves over the course of the MD. The analysis also guides the analysis of the helicoidal parameters of the most representative structure using CURVES.

Figures 5.4, 5.5, and 5.6 show intra-base pair, inter-base pair, and base pair-axis parameters obtained from Dials & Windows of the unmodified and modified CG decamer in the B form over 2 ns trajectories (except for CG^{8MMPh} that has a 4 ns trajectory), respectively, for the CG, CG^{8Ph}, CG^{8Tol}, CG^{8HMPh}, CG^{8CPh} and CG^{8MMPh} oligonucleotides. There was no significant deviation of the intra base pair parameters of C8-aryl modified CG decamer observed when compared to the unmodified DNA with the possible exception of STG and BKL (these are low energy processes). Minor perturbations in these two parameters were seen for CG^{8Ph}, CG^{8Tol} and perhaps CG^{8MMPh} at or near the modified base.

In the case of inter-base pair parameters (Figure 5.5), RIS and SLD of G6:C15 base pairs were altered. Both showed significant changes for CG^{8Ph}, CG^{8Tol}, and CG^{8HMPh}, while mainly RIS was different for CG^{8CPh} and CG^{8MMPh} containing oligonucleotides in the B-DNA form. The increase in rise parameter and the change in slide to negative value were likely driven by the tendency toward stacking of the aryl adduct of G6 and pyrimidine C5. Paralleling these differences is the effect of the C8-

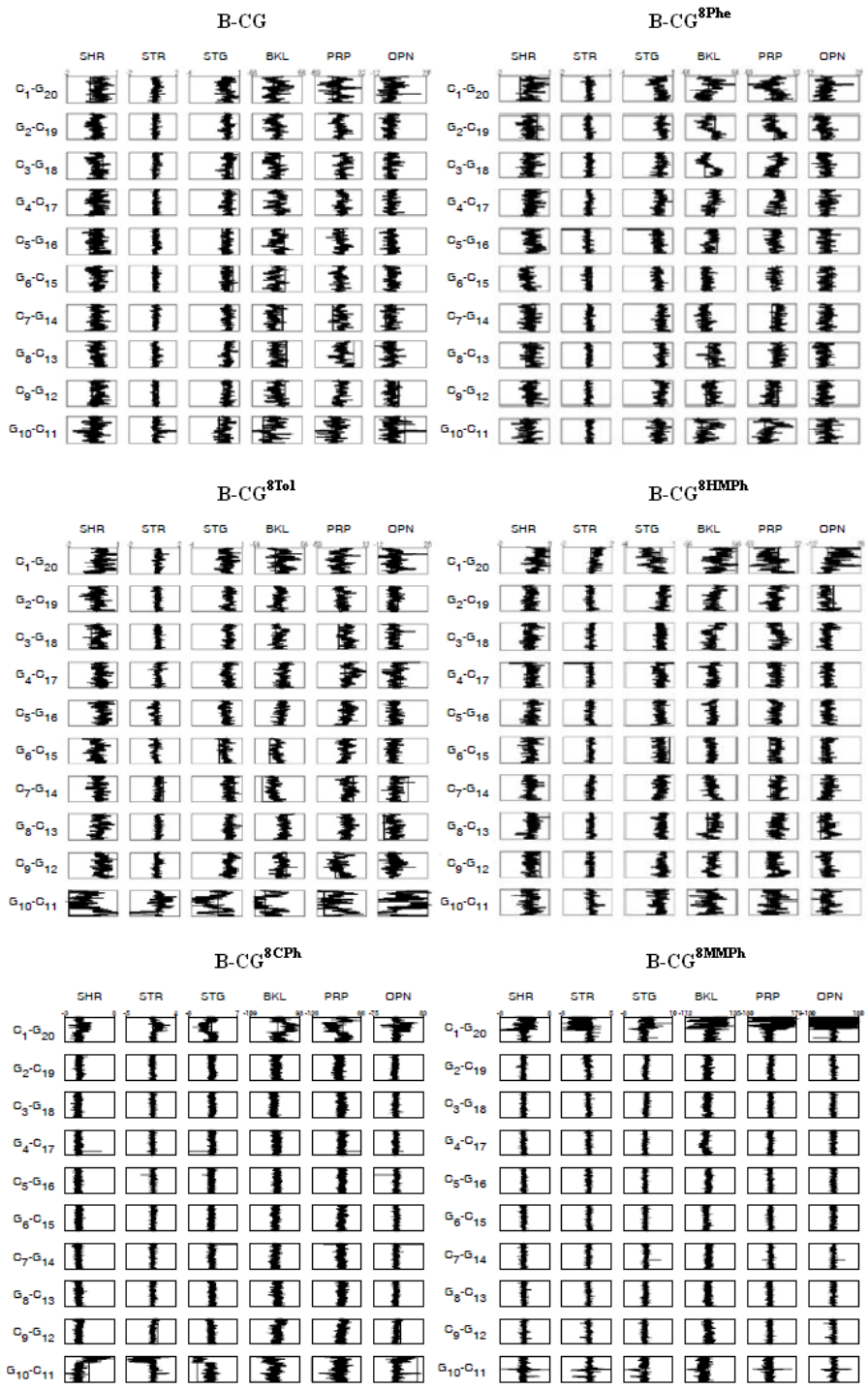


Figure 5.4 Intra-base pair parameters for all base pairs obtained from Dials & Windows analysis of the production trajectories of the unmodified and modified CG decamers in the B-DNA form.

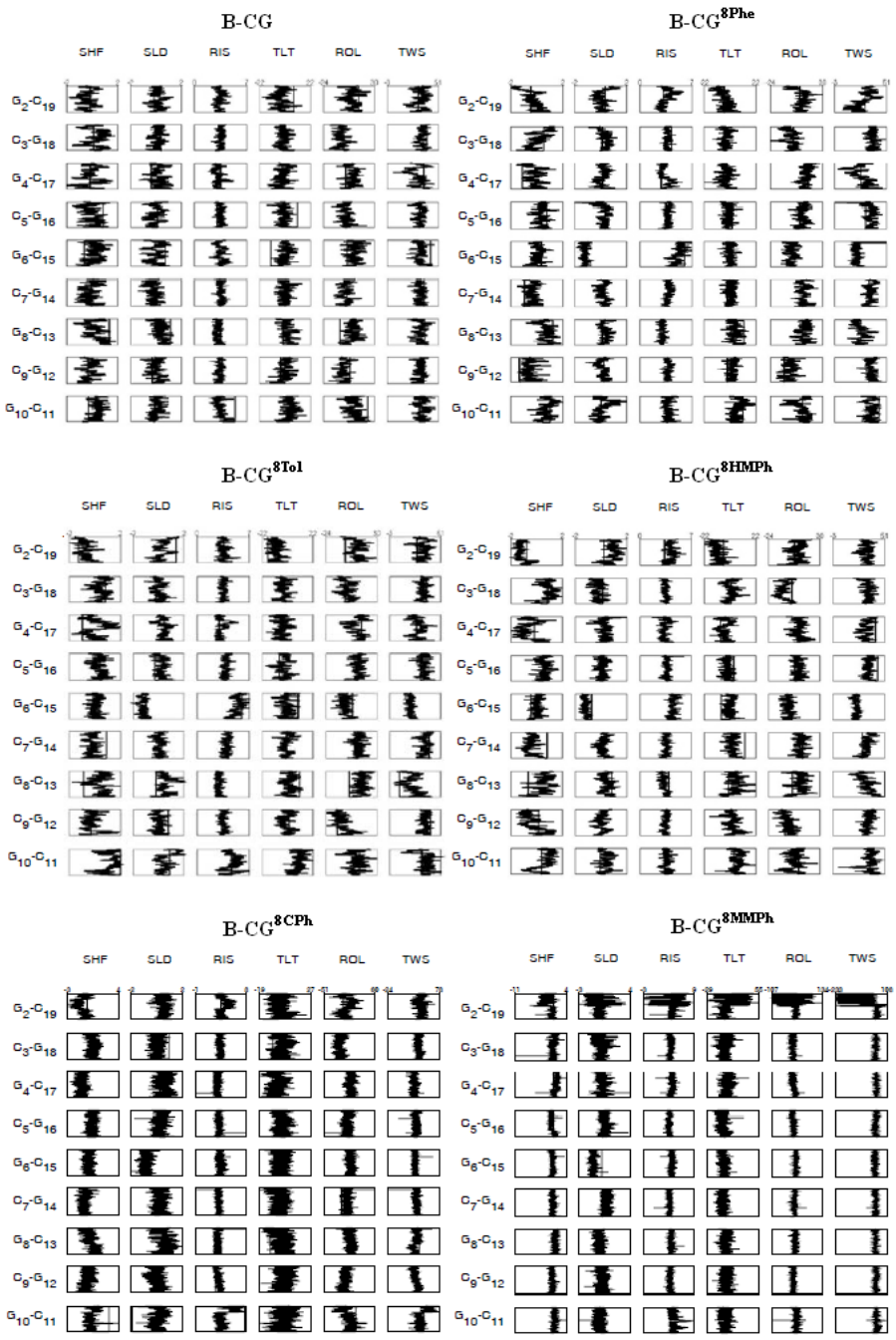


Figure 5.5 Inter-base pair parameters of all base pairs obtained from Dials & Windows analysis of the production trajectories of the unmodified and modified CG decamers in the B-DNA form.

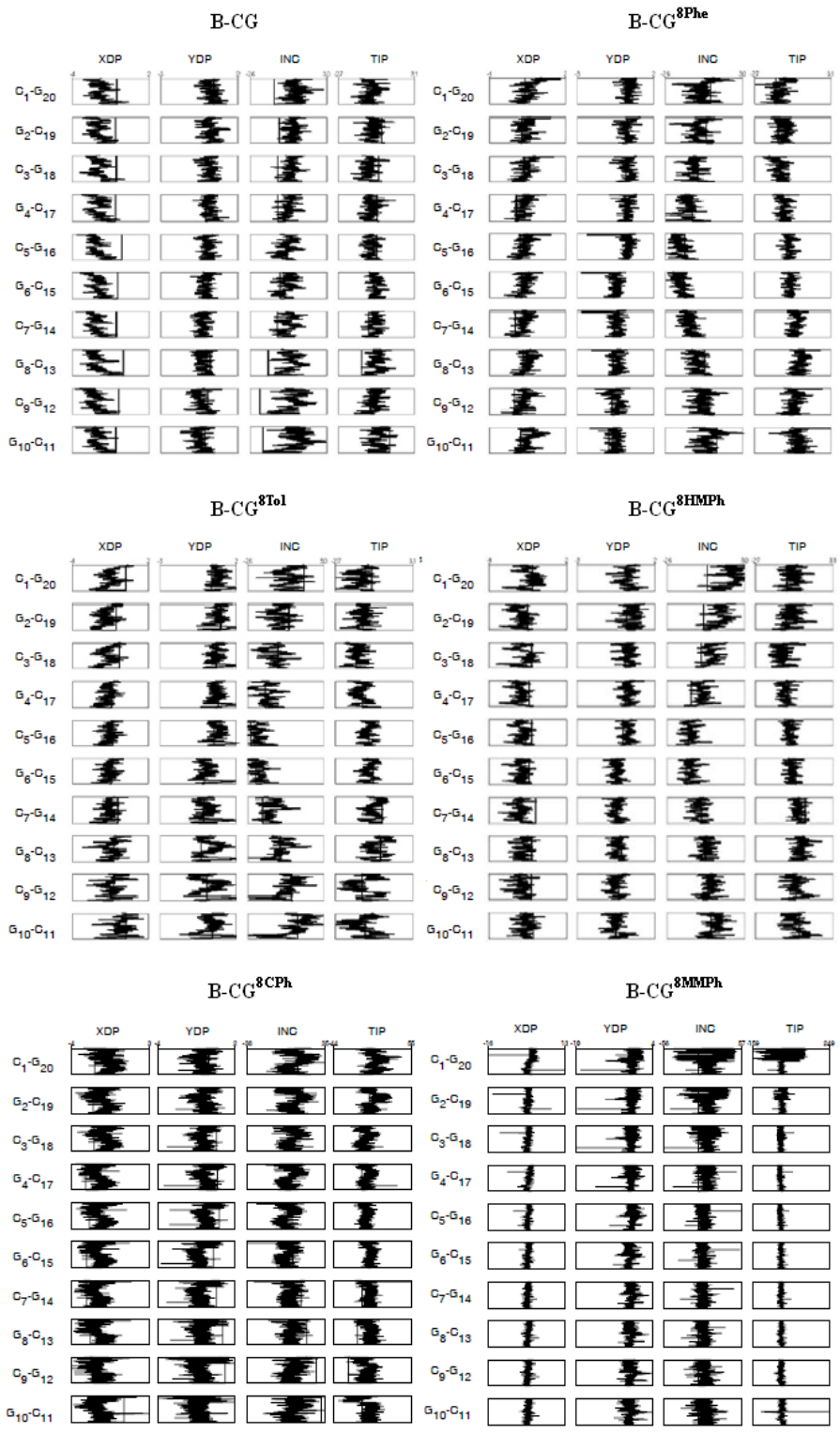


Figure 5.6 Base pair-axis parameters of base pairs obtained from Dials & Windows analysis of the production trajectories of the unmodified and modified CG decamers in the B-DNA form.

substituent on YDP and INC (and to a lesser extent, TIP) as seen in Figure 5.6. Thus, the oligonucleotides that displayed significant effects on both SLD and TLT display large changes in YDP and INC and had little effect on SLD or TLT, only YDP was altered (e.g., CG^{8CPh} and CG^{8MMPh}).

To put the changes pointed by Dials & Windows on a more quantitiative analysis, the helicoidal parameters were obtained utilizing the most representative structure. CURVES was used for this purpose with the results being compiled in Tables 5.1 and 5.2. Inspection of Table 5.1 indicates that STG and BKL are, as suggested by Dials & Windows, significantly affected by the C8-aryl adduct. It should be noted, however, that STG and BKL are not energetically costly motions and are more likely in response to other conformational changes.

As suggested by Dials & Windows, the inter-base parameters SLD and RIS are significantly different from the unmodified oligonucleotide. In all cases, SLD is nearly 1 Å more negative than the unmodified oligonucleotide. RIS varied from a modest increase of 0.57 Å for CG^{8CPh} to 3.12 Å for CG^{8Tol}. To a lesser extent, changes in SHF and TWS are seen. In the case of SHF, the change was smallest for the CG^{8Ph} (-0.49 Å) and largest for CG^{8MMPh} (-1.08 Å). The changes in TWS also range from showing a modest decrease for CG^{8MMPh} (3.15°) to very large decrease for CG^{8Tol} (12.14°) indicating the C8-arylguanine modification produces local unwinding of the helix.

Based on values for the helicoidal parameters obtained from CURVES, the amount of rise (RIS), slide (SLD), and shift (SHF) were all significantly altered by the presence of the C8-arylguanine base. The only difference between the aryl groups, however, is the *p*-substituent. Furthermore, the observed changes in RIS, SLD, and SHF

Table 5.1 Helical parameters of the unmodified and modified CG decamers determined by CURVES analysis of the most representative structures in the B-DNA form.[a]

Intra-Base Pair						
	CG	**CG^{8Ph}**	**CG^{8Tol}**	**CG^{8CPh}**	**CG^{8MMPh}**	**CG^{8HMPh}**
	G6/G16	G6/G16	G6/G16	G6/G16	G6/G16	G6/G16
SHR	-0.12 (-0.10)	-0.04 (-0.06)	0.01 (-0.27)	-0.08 (0.09)	0.40 (-0.30)	0.00 (0.20)
STR	-0.03 (0.00)	-0.20 (0.16)	-0.17 (0.10)	0.23 (0.18)	0.04 (-0.25)	-0.06 (0.06)
STG	0.53 (0.10)	-0.48 (-0.13)	-0.67 (-0.20)	-0.11 (0.09)	-0.30 (-0.51)	-0.33 (-0.08)
BKL	9.80 (1.10)	-0.93 (-3.27)	3.60 (-3.50)	-4.57 (-0.36)	-5.00 (-3.01)	-15.6 (-2.89)
PRP	-12.1 (-12.0)	-9.50 (-9.57)	-6.71 (-10.5)	-11.39 (-12.92)	-9.85 (-8.15)	-9.50 (-11.2)
OPN	-2.47 (-1.83)	4.90 (1.17)	-2.12 (4.43)	-0.04 (-2.31)	0.54 (-1.55)	2.09 (2.10)

Inter-Base Pair						
	CG	**CG^{8Ph}**	**CG^{8Tol}**	**CG^{8CPh}**	**CG^{8MMPh}**	**CG^{8HMPh}**
	C5:G6	C5:G6	C5:G6	C5:G6	C5:G6	C5:G6
SHF	0.81 (0.06)	0.32 (0.00)	0.36 (0.07)	-0.08 (-0.01)	-0.27 (-0.13)	-0.22 (-0.05)
SLD	-0.44 (-0.14)	-1.76 (-0.06)	-1.69 (0.11)	-1.27 (-0.01)	-1.26 (-0.06)	-1.63 (-0.04)
RIS	3.00 (3.18)	5.55 (3.58)	6.12 (3.83)	3.57 (3.38)	3.95 (3.89)	4.74 (3.64)
TLT	-5.60 (-1.70)	-0.65 (1.97)	-10.3 (1.40)	3.24 (0.00)	7.53 (6.75)	5.30 (0.11)
ROL	7.97 (5.04)	0.28 (7.24)	-6.00 (7.09)	8.62 (2.26)	5.14 (1.50)	6.86 (7.79)
TWS	32.2 (31.45)	21.5 (30.1)	20.6 (30.4)	33.05 (34.33)	29.05 (35.0)	22.0 (30.4)

Base Pair-Axis						
	CG	**CG^{8Ph}**	**CG^{8Tol}**	**CG^{8CPh}**	**CG^{8MMPh}**	**CG^{8HMPh}**
	G6/G16	G6/G16	G6/G16	G6/G16	G6/G16	G6/G16
XDP	-2.22 (-2.06)	-1.32 (-1.34)	-0.75 (-0.71)	-1.06 (-0.97)	-0.79 (-0.61)	-0.08 (0.08)
YDP	0.31 (-0.17)	-0.43 (-0.42)	-0.36 (-0.37)	-0.31 (-0.28)	-1.40 (-1.57)	-0.10 (-0.02)
INC	1.65 (3.98)	-11.5 (-4.97)	-17.9 (-8.50)	-1.57 (0.55)	-2.35 (-1.06)	-15.8 (-7.30)
TIP	-2.00 (-2.72)	1.00 (1.52)	2.50 (3.06)	2.29 (1.92)	11.99 (11.08)	1.61 (1.03)

[a] Values in parenthesis are for the average over all bases

all effect the extent to which the C8-aryl group of the modified base stacks over the 5′-cytosine. The CG^{8Ph} displays the most negative slide (-1.76) (the effect being augmented by STR (-0.20)) and gives rise to a significant stacking interaction between phenyl adduct and pyrimidine on the 5′ side. To accommodate these changes and relieve steric interactions, an increase in local rise (5.55 Å) occurs. The latter increase is responsible for the observed decrease in TWS. Note that while the observed changes result in the

stacking interaction and though similar SHF and SLD values are observed for the CG^{8Tol} derivative, the unusually large negative roll (-6.00) and increase in rise prevent the *p*-tolyl from stacking well over the 5′ pyrimidine.

The stacking interaction of CG^{8CPh}, CG^{8MMPh}, and CG^{8HMPh} were expected to be lower than CG^{8Ph} based on rise, slide, shift, and stretch parameters and is what is suggested by the most representative structures (Figure 5.2). In case of CG^{8HMPh}, there is a possibility that hydroxyl proton of *p*-CH_2OH H-bonding with nearby phosphate oxygen and this, in turn shifts slide back toward the unmodified oligonucleotide at the G6:C15 base pair. If stacking arrangement observed computationally (and experimentally) for the CG^{8Ph}, was adopted by CG^{8HMPh}, it would disrupt the optimum distance (2.4 Å) between hydroxyl proton and phosphate oxygen and diminish the chance of H-bonding.

The unfavorable electrostatic interaction between the negatively charge phosphate backbone of the oligonucleotide and the *p*-COO^- substituent in CG^{8CPh} appears to be the main reason for the decreased stacking in this oligonucleotide as the more interaction there is between the C5 pyrimidine and the aryl group of CG^{8CPh} the more the carboxylate group interacts with the negatively charged oxygens on the backbone. In addition, this oligonucleotide shows a fair amount of SHF, which further decreases stacking interactions.

The steric effect of the *p*-CH_2OCH_3 (with the phosphate backbone) is likely to be the cause of the decrease in stacking for $CG8^{MMPh}$ as it will decrease the amount of SLD relative to the unmodified oligonucleotide. Furthermore, the near zero values for inclination, a base pair-axis parameter, of G6:C15 of CG^{8CPh} (-1.57) and CG^{8MMPh} (-2.35) probably occurs to avoid unfavorable interactions between the *p*-substituents and the

backbone. Having the same degree of inclination like CG^{8Ph} (-11.5) would put *p*-COO^- and *p*-CH_2OCH_3 very close to the phosphate group of C5.

In addition to the effect of aryl adducts on base pair parameters that may contribute to B-DNA destabilization, groove parameters also affected by the presence of aryl adducts. The minor groove widths obtained from the CURVES analysis of the most representative structures are shown in Table 5.2. The C8-aryl adducts on G6 generally reduce the minor groove width which may result in increasing the electrostatic repulsion caused by bringing the phosphates close together. Although the trend of minor groove reduction and the B-Z conversion were not well correlated, the effect likely plays a role in reducing the stability of the B-DNA form of the C8-arylguanine modified oligonucleotides.

Table 5.2 Major and minor groove widths of the unmodified and modified CG decamers determined by CURVES analysis of the most representative structures in the B-DNA form.

Major Groove Width (A°)						
Base No.	**CG**	**CG^{8Ph}**	**CG^{8Tol}**	**CG^{8CPh}**	**CG^{8MMPh}**	**CG^{8HMPh}**
4	12.20	16.80	21.94	12.14	13.21	15.96
5	11.85	17.50	23.72	12.06	13.09	16.67
6	11.36	17.55	23.90	13.14	11.18	15.45
7	15.38	15.72	15.87	12.33	12.67	13.20
Average	12.70	16.89	21.36	12.42	12.54	15.32

Minor Groove Width (A°)						
Base No.	**CG**	**CG^{8Ph}**	**CG^{8Tol}**	**CG^{8CPh}**	**CG^{8MMPh}**	**CG^{8HMPh}**
4	7.91	7.22	5.00	6.60	5.73	6.83
5	8.40	4.32	4.81	4.65	6.51	6.38
6	7.00	5.99	5.41	7.01	5.47	7.74
7	7.77	7.58	8.80	9.40	5.74	7.13
Average	7.77	6.28	6.01	6.92	5.86	7.02

5.6.2 Z-DNA Oligonucleotides

The intra-base pair, inter-base pair, and base pair-axis parameters obtained by Dials & Windows analysis of the unmodified and modified CG decamers in Z form over 2-4 ns production runs are shown in Figure 5.7, 5.8, and 5.9. There was no significant difference observed between the unmodified and modified oligonucletides when the intra-base parameters (Figure 5.7) are considered except in the case of the CG^{8MMPh} derivative where greater variation in STR and STG are observed, suggesting greater mobility of this base and therefore decreased stability of the base-pair. Much the same can be said for the inter-base parameters (Figure 5.8). The alteration seen in SLD, RIS, ROL, and TWS are due to the alteration of the glycosidic bond of the pyrimidines (*anti*) and the purines (*syn*) in Z DNA. The main outlier is the SLD and RIS values for the C6-G15 base pair. In most cases, SLD becomes more positive while RIS becomes more negative while in the case of the CG^{8CPh} and CG^{8MMPh}, just the opposite is observed. Finally, no real differences are observed for the base pair-axis parameters. The different scales used for the last two panels corresponding to CG^{8CPh} and CG^{8MMPh} seem to suggest a difference but this is only a scaling effect.

The helical parameters from the CURVES analysis of the most representative Z-DNA structures are shown in Table 5.3. Since the C8-aryl group of the modified guanine lies largely outside of the DNA helix in the Z-form, the aryl group should have little effect on DNA conformation (constraining the view to just the Z-DNA form). For the most part, and as expected from the most representative structures of CG^{8Ph}, CG^{8Tol}, and CG^{8HMPh} in Z-DNA form shown in Figure 5.3, there were not many differences between the modified and unmodified CG oligonucleotides. The most notable differences seen in

the intra-base pair parameters are 1) the value for STR of CG^{8MMPh}, which is quite negative, 2) PRP for both CG^{8CPh} and CG^{8MMPh}, both of which departed considerably from the unmodified oligonucleotide, albeit in opposite directions and 3) OPN for the CG^{8MMPh} oligonucleotide.

Differences were also observed in inter-base pair helical parameters. The overall trends are that the CG^{8CPh} and CG^{8MMPh} modified oligos alter inter-base parameters typically in the opposite direction as compared to the CG^{8Ph}, CG^{8Tol} or CG^{8HMPh} oligonucleotides (with the unmodified oligo as reference). The deviations from the unmodified oligonucleotide can be quite large (selected values of RIS, TLT, and ROL).

In the case of Z-CG^{8CPh}, an obvious outcome (see Figure 5.3) is that the aryl adduct is almost planar to purine ring of G6 unlike the rest of modified CG decamers in which the aryl group is approximately perpendicular to purine. It is not clear what the source of this unique structural feature of Z-CG^{8CPh} is though the carboxylate group is very electron donating and therefore there may be a resonance effect with the relatively electron deficient guanine base. Such an effect would also reduce the base pair strength and increase base mobility (in turn, increasing BKL, PRP, OPN, etc.). This proposal is partially supported by the Dials & Windows results that showed considerable variation of helical parameters in the MD trajectories of CG^{8MMPh}, especially in the middle of sequence at the site of modification.

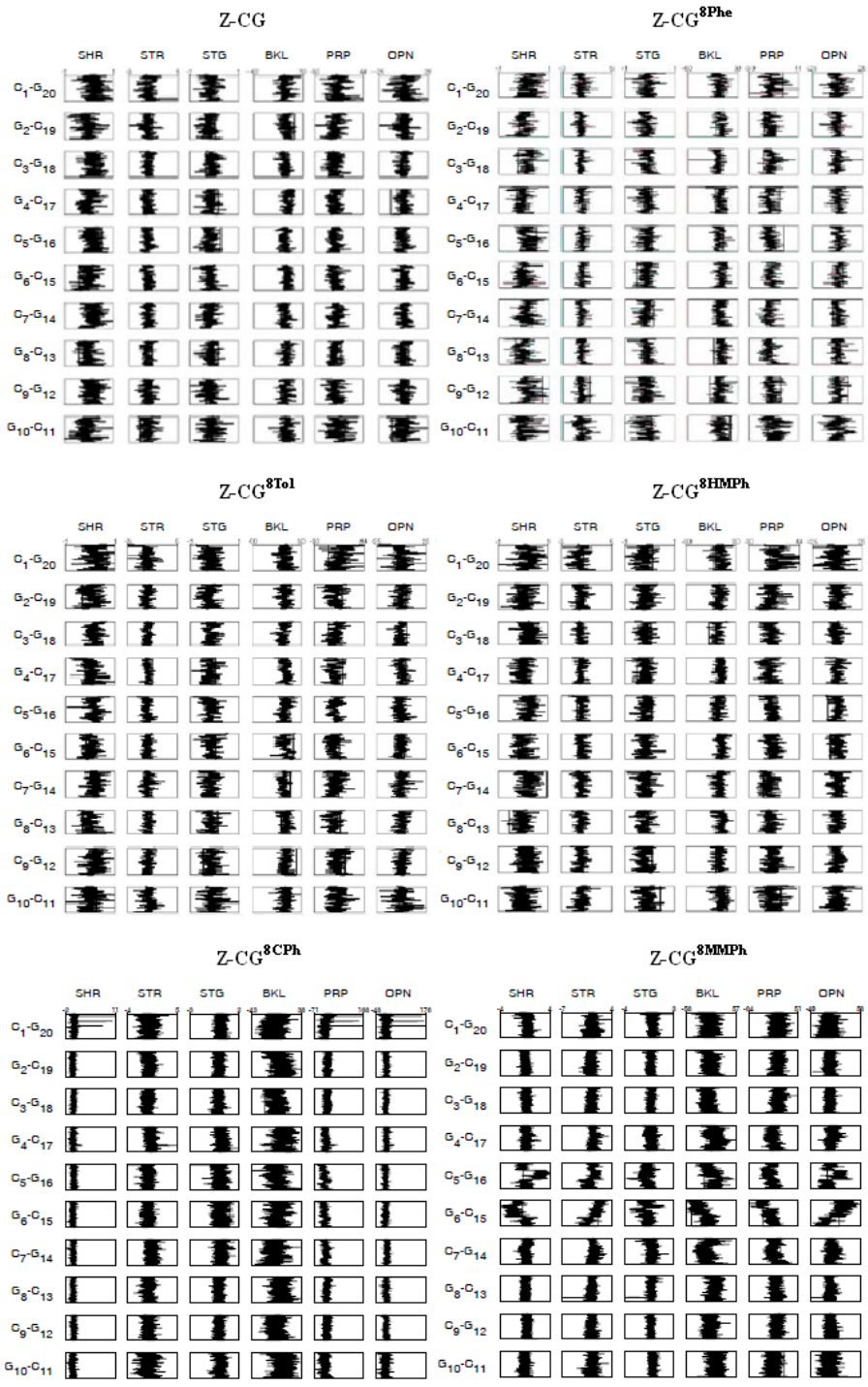


Figure 5.7 Intra-base pair parameters for all base pairs obtained from Dials & Windows analysis of the production trajectories of the unmodified and modified CG decamers in the Z form.

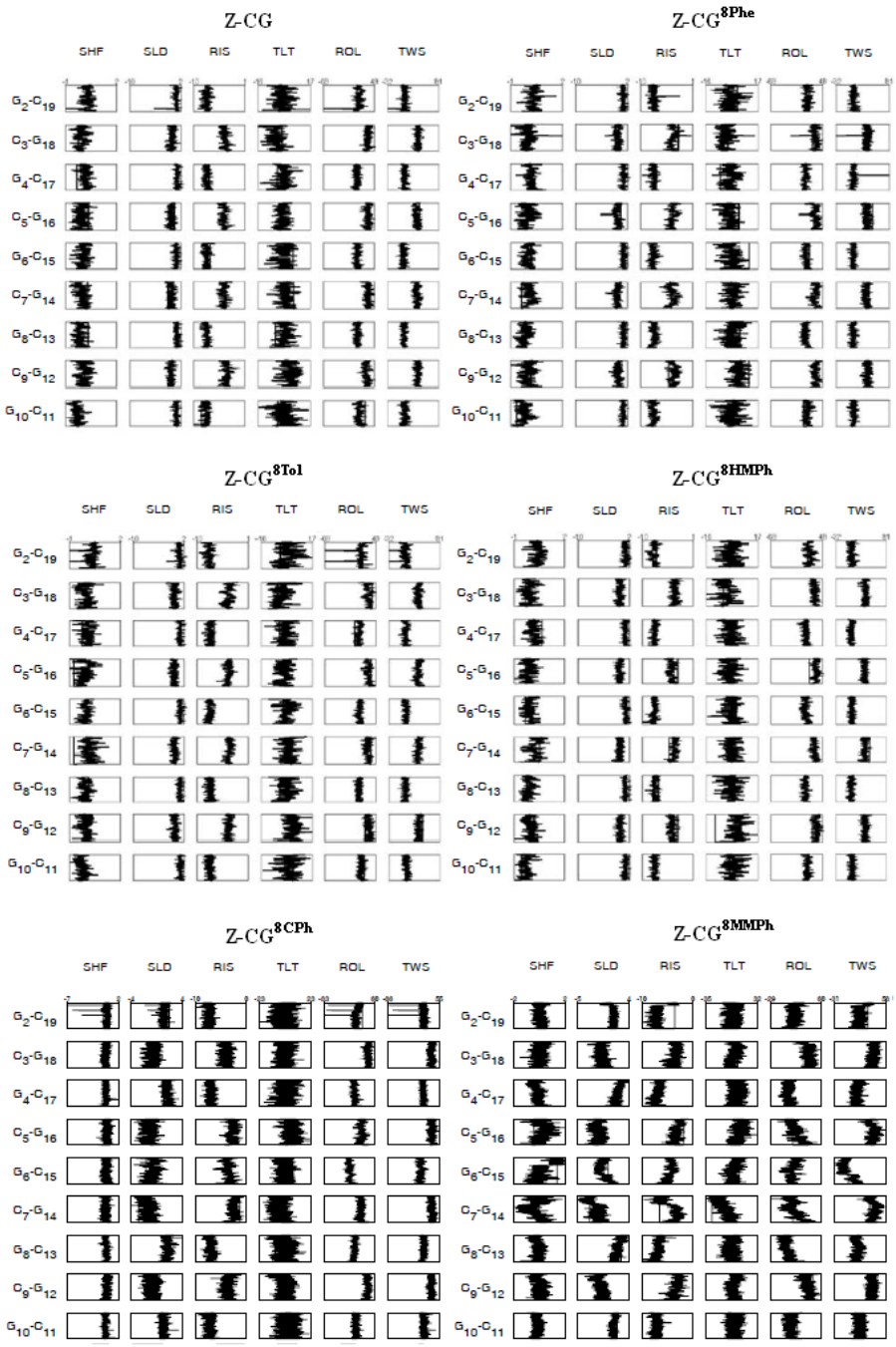


Figure 5.8 Inter-base pair parameters for all base pairs obtained from Dials & Windows analysis of the production trajectories of the unmodified and modified CG decamers in the Z form.

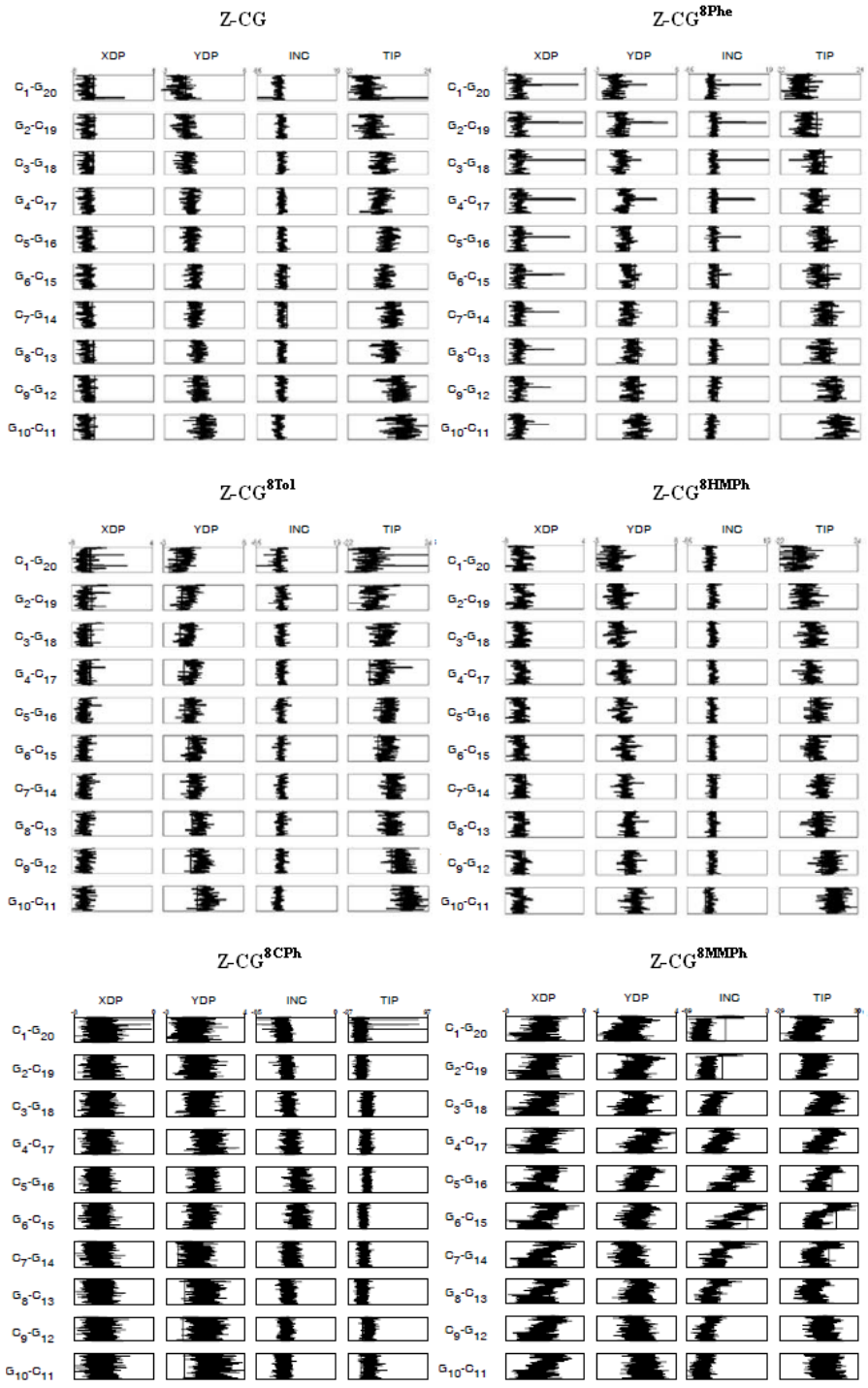


Figure 5.9 Base base-axis parameters for all base pairs obtained from Dials & Windows analysis of the production trajectories of the unmodified and modified CG decamers in the Z form.

With the exception of CG^{8CPh} and CG^{8MMPh} with structures seemed to be relatively less sound, the modified CG decamers in Z-form look more rigid and the aryl adducts did not appear to interfere with the conformation, in general, based on helical parameters. Interestingly, minor groove widths in Z-DNA were wider in case of the modified CG decamers except for CG^{8HMPh}. This effect is the reverse of what is observed for the B forms and could be one of the factors that stabilize Z form because increasing the minor groove width decreases the unfavorable electrostatic interaction.

Table 5.3 Helical parameters of the unmodified and modified CG decamers determined by CURVES analysis of the most representative structures in the Z-DNA form.[a]

Intra-Base Pair							
		CG	**CG^{8Ph}**	**CG^{8Tol}**	**CG^{8CPh}**	**CG^{8MMPh}**	**CG^{8HMPh}**
		G6/G16	G6/G16	G6/G16	G6/G16	G6/G16	G6/G16
SHR		-0.25 (-0.30)	0.03 (-0.02)	0.00 (-0.09)	-0.18 (-0.01)	-0.09 (0.04)	-0.04 (-0.09)
STR		0.03 (0.09)	0.24 (0.10)	-0.04 (0.08)	0.17 (0.04)	-0.75 (-0.06)	0.22 (0.08)
STG		0.42 (0.00)	0.20 (0.20)	0.10 (0.11)	-0.02 (0.14)	0.94 (0.29)	-0.20 (0.02)
BKL		-3.50 (-2.30)	-8.60 (1.07)	-2.50 (0.80)	9.09 (1.32)	14.18 (1.53)	-9.30 (-6.10)
PRP		-5.40 (0.56)	-2.00 (-2.70)	4.20 (4.00)	-25.28 (-11.37)	-34.74 (-13.11)	3.10 (1.20)
OPN		1.15 (-0.70)	0.45 (0.04)	0.44 (0.06)	4.01 (-0.38)	-19.60 (-2.34)	1.26 (-0.21)

Inter-Base Pair							
		CG	**CG^{8Ph}**	**CG^{8Tol}**	**CG^{8CPh}**	**CG^{8MMPh}**	**CG^{8HMPh}**
		C5:G6/ G6:C7	C5:G6/ G6:C7	C5:G6/ G6:C7	C5:G6/ G6:C7	C5:G6/ G6:C7	C5:G6/ G6:C7
SHF	(C:G)	0.19 (0.03)	0.44 (0.35)	0.27 (0.28)	-0.53 (0.06)	-0.39 (0.27)	0.21 (0.14)
	(G:C)	-0.44 (-0.93)	-0.49 (-0.42)	-0.22 (-0.19)	0.17 (0.07)	0.28 (-0.32)	-0.53 (-0.12)
SLD	(C:G)	4.09 (4.26)	4.37 (4.27)	4.00 (4.22)	1.84 (3.65)	1.57 (3.93)	4.35 (4.31)
	(G:C)	-4.48 (-4.17)	-4.67 (4.15)	-3.81 (-4.00)	-4.32 (-3.56)	-3.99 (-3.88)	-4.42 (-4.27)
RIS	(C:G)	3.91 (4.26)	3.57 (4.34)	3.92 (4.46)	4.81 (5.17)	2.88 (4.36)	4.15 (3.71)
	(G:C)	3.03 (2.86)	2.98 (2.88)	2.81 (2.77)	3.08 (2.71)	6.08 (3.79)	2.88 (3.35)
TLT	(C:G)	4.90 (-3.58)	8.53 (-2.78)	2.04 (-2.49)	-5.38 (3.19)	-1.78 (-1.86)	-8.28 (2.31)
	(G:C)	3.22 (4.55)	0.45 (2.65)	1.50 (0.18)	3.27 (-0.31)	13.60 (-0.55)	-3.63 (-1.37)
ROL	(C:G)	-1.90 (1.67)	7.07 (1.97)	11.47 (6.79)	-26.47 (-5.51)	11.61 (-0.62)	4.01 (4.18)
	(G:C)	3.73 (4.32)	-8.74 (0.55)	-2.76 (3.70)	3.41 (10.80)	-4.04 (4.99)	-7.90 (-2.28)
TWS	(C:G)	-8.49 (-10.7)	-10.6 (-8.78)	-10.70 (-13.4)	-4.99 (-9.73)	13.46 (-5.76)	-12.3 (-10.8)
	(G:C)	-44.7 (-45.3)	-44.4 (-44.9)	-38.6 (-41.6)	-35.75 (-41.11)	-39.29 (-42.90)	-48.1 (-43.2)

Base Pair-Axis							
		CG	**CG^{8Ph}**	**CG^{8Tol}**	**CG^{8CPh}**	**CG^{8MMPh}**	**CG^{8HMPh}**
		G6/G16	G6/G16	G6/G16	G6/G16	G6/G16	G6/G16
XDP		-2.15 (-2.18)	-2.00 (-2.03)	-1.55 (-1.55)	-0.21 (-0.47)	-1.72 (-1.64)	2.15 (-2.15)
YDP	(C:G)	-2.17 (-2.07)	-1.86 (-1.80)	-1.79 (-1.83)	-1.68 (-2.14)	-0.92 (-1.26)	-2.15 (-2.10)
	(G:C)	2.05 (2.18)	2.41 (2.35)	2.30 (2.34)	1.29 (1.47)	2.20 (2.71)	2.11 (2.11)
INC		6.28 (7.78)	5.80 (6.76)	11.0 (10.5)	11.66 (16.44)	-1.10 (7.42)	7.47 (6.33)
TIP		179 (178)	180 (181)	182 (18.2)	1 (35.6)	-171 (-139)	179 (179)

[a] Values in parenthesis are for the average over all bases

Table 5.4 Major and minor groove widths of the unmodified and modified CG decamers determined by CURVES analysis of the most representative structures in the Z-DNA form.

Major Groove Width (A°)						
Base No.	**CG**	**CG^{8Ph}**	**CG^{8Tol}**	**CG^{8CPh}**	**CG^{8MMPh}**	**CG^{8HMPh}**
4	12.28	10.92	13.27	14.48	22.91	12.92
5	14.41	14.36	15.01	15.70	20.69	15.60
6	10.50	12.40	11.76	15.29	19.42	11.15
7	N/A	N/A	15.31	N/A	N/A	N/A
Average	12.40	12.56	13.84	15.16	21.01	13.22

Minor Groove Width (A°)						
Base No.	**CG**	**CG^{8Ph}**	**CG^{8Tol}**	**CG^{8CPh}**	**CG^{8MMPh}**	**CG^{8HMPh}**
4	1.93	1.49	1.91	4.87	5.04	2.23
5	4.15	4.79	4.79	7.03	5.14	4.29
6	4.13	3.75	4.22	5.38	5.07	2.63
7	4.13	4.98	4.13	4.99	5.08	4.27
Average	3.59	3.75	3.76	5.57	5.08	3.36

5.7 Free Energy Calculation

The MM_PBSA module of Amber was used to calculate the absolute free energies of the modified DNAs in solution from Amber trajectories. The following set of equations show the breakdown of the components of the free energies obtained with MM_PBSA.

$$GAS = ELE + VDW + INT \qquad \text{Equation 5.1}$$

$$PBSOL = PBSUR + PBCAL \qquad \text{Equation 5.2}$$

$$PBELE = PBCAL + ELE \qquad \text{Equation 5.3}$$

$$PBTOT = PBSOL + GAS \qquad \text{Equation 5.4}$$

$$TSTOT = TSTRA + TSROT + TSVIB \qquad \text{Equation 5.5}$$

$$\Delta G\_PB = PBTOT - TSTOT \qquad \text{Equation 5.6}$$

From Equation 5.4 - 5.6, the free energy of the system (ΔG_PB) is calculated as the summation of the molecular mechanic energy (GAS), the solvation free energy (PBSOL), and the total entropy multiplied by the temperature (TSTOT). Sander was used to determine the GAS term (Equation 5.1), which represents electrostatic interaction (ELE), van der Waals (VDW), and internal energy (INT). Poisson-Boltzman (PB) approach using Delphi[102] was applied to calculate the hydrophobic contribution to solvation free energy (PBSUR) and the reaction field energy (PBCAL) which are summed to give PBSOL (Equation 5.2). The total entropy (TSTOT), composed of translational (TSTRA), rotational (TSROT), and vibrational (TSVIB) entropies, were calculated using the nmode module in Amber

The production trajectories from frame 500 to 1950 ps were used in the MM_PBSA calculation after water and ions were stripped out. To insure the MM_PBSA results for CG^{8CPh} and CG^{8MMPh} were consistent and comparable with the previous data set for CG, CG^{8Ph}, CG^{8Tol}, and CG^{8HMPh}, MM_PBSA in Amber version 7[110] was used. Free energy calculations were performed on both the B and Z forms and by subtracting the free energy of B-DNA (ΔG_PB_B) from the free energy of Z-DNA (ΔG_PB_Z), ΔΔG_PB_ZB was obtained. This value was then used to compare the effect of aryl adducts on B-Z conversion and determine the relative stability of the two forms. A negative value for ΔΔG_PB_ZB indicates that the transition from B to Z is energetically favorable. The free energies of the modified and unmodified CG decamers in B and Z form calculated by MM_PBSA method are reported in Table 5.5.

Sietraj was used as an alternative method to MM_PBSA to calculate the free energy from molecular dynamic trajectory of CG decamers. Rather than setting DNA as a

single molecule in the system and calculating the absolute free energy, as we did in MM_PBSA calculation, a DNA duplex was considered to be two molecules binding, one a receptor and the second a ligand, though these two parts are identical. The binding free energy (ΔG) was then calculated as the difference in free energies between the isolated palindrome strands that comprise the duplex and the duplex itself. Generally, sietraj uses a similar approach as MM_PBSA to obtained ΔG with the following exceptions: sietraj neglects the vibration entropy, a different surface generation method is used[107], and sietraj uses an internal Poisson solver[111,112]. As with the MM_PBSA calculations, a ΔG_ZB was calculated. The binding free energies of the modified and unmodified CG decamers in both B and Z form are reported in Table 5.6 as is ΔG_ZB negative values indicate the Z-DNA form is preferred.

Table 5.5 The free energies of the unmodified and modified CG decamers calculated with the MM_PBSA method

Energy (kcal/mol)	CG		CG^{8Ph}		CG^{8Tol}		CG^{8HMPh}		CG^{8CPh}		CG^{8MMPh}	
	B	Z	B	Z	B	Z	B	Z	B	Z	B	Z
ELE	-817.40	-63.41	-821.80	-65.69	-698.30	53.67	-812.71	-24.78	-765.59	-360.68	-730.40	-470.84
VDW	-164.30	-196.20	-169.60	-194.20	-171.80	-195.20	-171.10	-196.50	-177.38	-161.64	-153.52	-147.59
INT	939.70	953.50	968.40	975.80	975.20	981.00	955.90	983.50	1111.28	1161.11	1150.98	1244.11
GAS	-42.00	693.89	-23.00	715.91	105.10	839.47	-27.91	762.22	168.31	638.79	267.06	625.68
PBSUR	17.63	16.43	17.90	17.14	17.99	17.35	17.89	17.43	18.37	17.48	30.80	30.72
PBCAL	-4858.00	-5604.00	-4839.00	-5593.00	-4977.00	-5729.00	-4838.00	-5625.00	-4635.08	-5030.51	-4778.74	-5036.44
PBSOL	-4841.00	-5587.00	-4821.00	-5576.00	-4959.00	-5712.00	-4820.00	-5607.00	-4616.72	-5013.03	-4747.94	-5005.72
PBELE	-5676.00	-5667.00	-5661.00	-5659.00	-5676.00	-5675.00	-5651.00	-5649.00	-5400.67	-5391.19	5509.13	-5507.29
PBTOT	-4883.00	-4893.11	-4844.00	-4860.09	-4853.90	-4872.53	-4847.91	-4844.78	-4448.41	-4374.24	-4480.88	-4380.04
TSTRA	15.57	15.57	15.59	15.59	15.60	15.60	15.60	15.60	15.61	15.61	15.61	15.61
TSROT	15.18	15.20	15.22	15.23	15.23	15.24	15.21	15.25	15.19	15.29	15.24	15.32
TSVIB	490.30	478.70	497.30	493.30	501.70	499.30	500.30	503.00	526.67	521.01	533.88	533.15
TSTOT	521.05	509.47	528.11	524.12	532.53	530.14	531.11	533.85	557.47	551.91	564.73	564.08
ΔG_PB	-5404.05	-5402.58	-5372.11	-5384.21	-5386.43	-5402.67	-5379.02	-5378.63	-5005.88	-4926.15	-5045.61	-4944.12
ΔΔG_PB_ZB	1.47		-12.10		-16.24		0.39		79.73		101.49	

Table 5.6 The free energies of the unmodified and modified CG decamers calculated with the sietraj method.

Energy (kcal/mol)	CG		CG^{8Ph}		CG^{8Tol}		CG^{8HMPh}		CG^{8CPh}		CG^{8MMPh}	
	B	Z	B	Z	B	Z	B	Z	B	Z	B	Z
ΔG	673.71	649.39	676.46	646.85	663.48	649.40	656.04	634.13	851.66	782.18	780.43	740.72
ΔΔG_ZB	-24.32		-29.61		-14.08		-21.91		-69.48		-39.71	

The experimental results obtained by CD and NMR have shown that the order of Z-DNA preference is CG^{8CPh}, followed by CG^{8Ph}, CG^{8MMPh}, CG^{8HMPh}, then CG^{8Tol}, and finally CG based on the required salt concentration to obtain 50% of Z-DNA form. The free energy analysis study has been conducted in order to try to elucidate the factors underlying these results and therefore be able to explain the effect of selected aryl adducts on both B and Z-DNA stability and help us understand how the aryl adducts cause the conformational change.

The destabilization effect of aryl adduct on the B-DNA conformation was expected. This is based on the well known effect of the steric interaction that occurs between the the C8-arylguanine and the H-2′ proton which can be avoided if the nucleoside adopts a *syn* conformation. However, for base pair formation, the C8-arylguanine base must adopt the *anti* in B-DNA. From the structural analysis of the computational data, B-DNA destabilization effect of C8-arylguanine adducts can be inferred from the helicoidal parameters that show the modified oligonucleotides are distorted from the unmodified oligonucleotide when in the B-DNA form. Besides steric effects, structure analysis on different *p*-substituent derivatives has suggested that an electronic effect may also be important in the B-DNA destabilization as seen in case of CG^{8CPh}. The dual interactions, steric and electronic, of *p*-carboxyphenyl adduct may explain its remarkable effect on B-Z conversion.

A more interesting question is how the aryl adducts affect Z-DNA stability. Because the aryl adducts can drive B to Z form without significantly increasing denaturation, which could be a reasonable alternative to B-Z conversion, one would expect that the C8-aryl adducts should stabilize Z form or at least would not destabilize it.

Structural analysis has shown the similarity of helical parameters among the unmodified CG, CG^{8Ph}, CG^{8Tol}, and CG^{8HMPh} suggesting that phenyl, *p*-tolyl, and *p*-hydroxymethylphenyl did not alter the conformation in any significant extent. On the other hand, *p*-carboxyphenyl and *p*-methoxymethylphenyl have distorted the Z form of CG decamer in the way that would destabilize Z form which would contradict to our initial idea toward these two adducts. In contrast, since, in the Z-DNA form, the purines adopt a *syn* conformation and substituents located on the C8-guanine position lie outside of the DNA helix, little effect upon the stability of the Z-DNA forms is expected. This may be modulated some by the substituents attached to the phenyl ring. Groups that are hydrophobic will tend to destabilize the Z-DNA form while just the opposite is expected from hydrophilic substituents.

The free energy analysis of the models, then, should help verify the proposed effects and also determine the relative weights of steric and electronic interactions. MM_PBSA and sietraj were used to calculate the free energies of the unmodified and modified CG decamers in both B and Z form. As seen in Table 5.5, MM_PBSA calculation predicts that B-Z conversion would be favorable if CG decamer is modified with a phenyl or tolyl adduct. Based on ΔG_PB_ZB, the remaining adducts should prefer to adopt the B-DNA form, opposite from what we have observed experimentally. The results from sietraj are different than the MM_PBSA and predict that Z-DNA should be preferred for all of the CG decamers. It could be that the 'zero' point needs to be calibrated (e.g., set the free energy of the unmodified decamer to favor the B form). However, the predicted stability trend is also out of order based on the experimental results.

While the overall free energy values do not correlate with the experimental trends, the trends of the individual components, in many cases, make some sense. Shown in Figure 5.10 are plots of GAS and PBSOL. The unmodified and the modified oligonucleotides are arranged along the X-axis with respect to the salt concentration at which the f_Z is 0.5. The Y-axis is the indicated value calculated by MM_PBSA. The plot for GAS shows a slight downward trend in the stability of the B forms with the methoxymethyl and carboxy derivative calculated to be least stable ones. The B forms of CG^{8HMPh} and CG^{8Ph} are less destabilized than might be predicted assuming a linear relationship, but the previously discussed hydrogen bonding interactions form *p*-CH_2OH and stacking interactions for phenyl must attenuate the unfavorable steric interactions. In contrast, relative to the unmodified oligonucleotide, the modified base can either destabilize the Z form (CG^{8Tol}), have little effect (CG^{8HMPh} or CG^{8Ph}), or stabilize it (CG^{8CPh} or CG^{8MMPh}). The effect may, in part, be due to the hydrophobicity/hydrophilicity of the aryl group. For example, the tolyl group is the most hydrophobic and the carboxy the most hydrophilic which parallels the Z destabilization or stabilization caused by these groups relative to the unmodified derivative.

The plot for PBSOL, for the B form is fairly consistent with the expections. Ingoring the values for the CG^{8CPh} (these are the ones that are at 100 mM on the X-axis), there is little variation in PBSOL. PBSOL is the sum of PBSUR and PBCAL (the reaction field term) and is dominated by the latter. The specific values PBSOL for CG^{8CPh} (B or Z forms) suggest it is destabilized. As previously discuss, we expect the *p*-carboxylate group to destabilize the B-DNA form and may also destabilizing the Z-DNA form due to the increase in negative charge, the factor that most destabilizes the Z form

of DNA. Then, if the numbers that are used to calculate ΔG generally make sense, then why aren't the computational results better correlated with the experimental results? In part the problem is that the numbers being used to calculate the difference in free energy between the B and Z forms are both large but the difference is small. Thus, small errors in the free energy calculation for the B form or Z form are much less significant than when the difference is calculated.

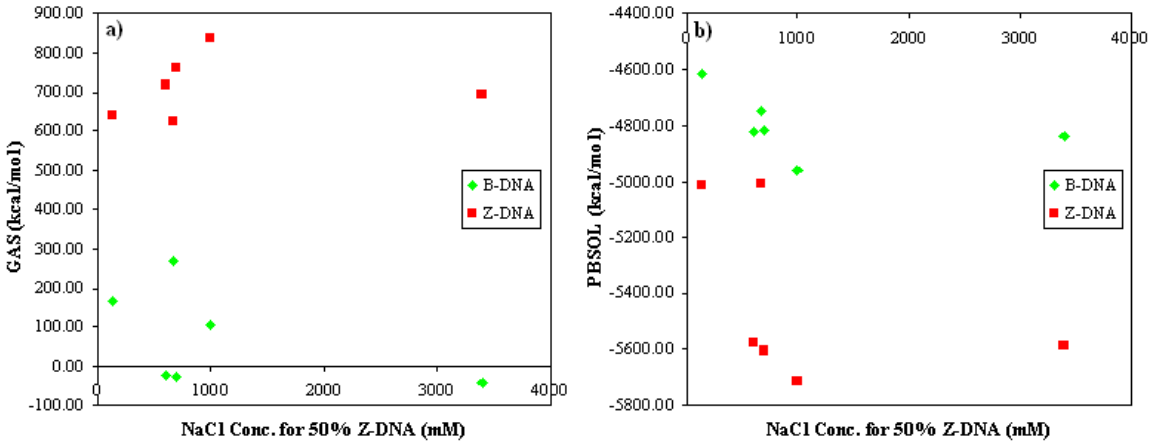


Figure 5.10 Plots of a) GAS or b) PBSOL vs NaCl concentration required to have 50% Z conformation of CG (3600 mM), CG^{8Tol} (1000 mM), CG^{8HMPh} (694 mM), CG^{8MMPh} (669 mM), and CG^{8CPh} (131 mM).

The current data from the free energy calculations from both methods are inconclusive and cannot be used to accurately predict preferred conformational stabilities (B or Z). However, this does not appear to be due to a fundamental problem with the approach rather it more like is a problem with the accuracy of the method used. Refinement of the methods, may in time, allow more accurate predictions.

CHAPTER 6
DISCUSSION AND CONCLUSIONS

Aryl hydrazines and related compounds are known for their carcinogenicity[4]. The mechanism of carcinogenesis is believed to initiate from metabolic activation[5,6] of aryl hydrazines which lead to the formation of harmful reactive intermediates, arenediazonium ions and aryl radicals. Consequently, DNA adducts can form from the reactive aryl hydrazine metabolites and these have been suggested to be the cause of genetic alterations which eventually lead to carcinogenesis. Among several potential hypotheses, carcinogenesis through Z-DNA stabilization by the DNA aryl adduct has particularly caught our attention and has been one of our major research questions.

To evaluate the relevant of Z-DNA stability in aryl hydrazine carcinogenesis, the effect of various C8-aryl adducts on B-Z conversion has been investigated using both experimental and computational methods. We have reported on the CG decamers that contain a C8-phenylguanine modified base[22]. The modification forms from arenediazonium ions or aryl radicals generated during the metabolism of phenyl hydrazine and may be involved in the mutagenicity of phenyl hydrazine. The resulting adduct causes a shift in the B/Z-DNA equilibrium toward the Z-conformation. In this study, additional examples of CG decamers that contain C8-arylguanine adducts is described. The adducts selected for study form by metabolism of known carcinogenic aryl hydrazines. The correlation between aryl hydrazine carcinogenicity and the shift in B/Z-DNA equilibrium were expected, if Z-DNA stabilization or/and B-DNA destabilization is the only major factor involved in aryl hydrazine carcinogenesis.

The aryl modified dG phosphoramidites have been synthesized in order to prepare the modified CG decamers through automated DNA synthesis. CD and NMR analysis have been extensively used in this study to examine the effect of selected aryl adducts on Z-DNA stability. The experimental results have shown there is a remarkable effect caused by the C8-arylguanine adduct on B-Z equilibrium which is partially supported by the computational study. The results from both methods will be discussed in this chapter. The connection between aryl hydrazine carcinogenesis and Z-DNA stabilization/ B-DNA destabilization will also be addressed.

6.1 Modified Oligonucleotide Synthesis

Generally speaking, the preparation of the phosphoramidites used to prepare the C8-arylguanine modified oligonucleotides was the same as previously reported[22]. The syntheses utilized a Suzuki coupling to introduce the aryl residue, and the subsequent steps used to introduce the protecting groups, are standard. Nevertheless, a few adaptations were made, particularly for the preparation of the *p*-hydroxymethylphenyl derivative, which requires TBS protection of the benzyl alcohol group.

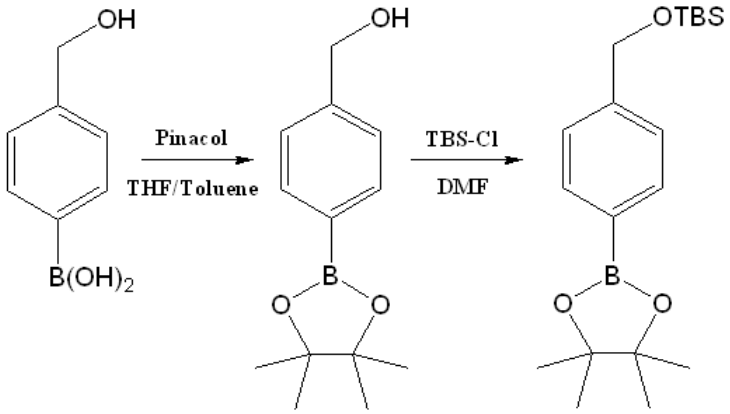
Using intermediates that contained the unprotected *p*-hydroxymethyl substituent was found to be problematic. The nucleophilic nature of the free hydroxyl group was expected to interfere with reactions that mechanistically required nucleophilic substitution. However, due to the success we had making the CG^{8CPh} oligonucleotide without using a protected *p*-carboxyphenyl dG phosphoramidite, the synthesis of CG^{8HMPh} without *p*-hydroxymethyl protection was attempted. Although, the free *p*-hydroxymethyl substituent did not hinder Suzuki coupling or N2 protection, 5′OH

protection was affected. Similar to 5′OH, the *p*-hydroxymethyl substituent seemed to also be reactive toward DMTr-Cl. Thus, we obtained a bis-tritylated product. This product was isolated and used to prepare the phosphoramidite. However, the use of this material for automated DNA synthesis failed to provide the desired oligonucleotide, CG^{8HMPh}. Failure to make the desirable oligonucleotide confirmed that protection of the *p*-hydroxymethyl substituent would be necessary.

The TBS group, which has commonly been used to protect the 2′-OH hydroxyl group of the ribose sugar in automated RNA synthesis[68], was selected to serve as the protecting group for *p*-hydroxymethyl substituent since it is known to survive automated synthesis and can be removed with the same reagent as used to remove other protecting groups following automated synthesis. Initially, 4-(TBS-O-methyl)-phenyl boronate pinacol ester was made (Scheme 6.1) and used in the in Suzuki coupling to make compound **(6)**. Likely due to the hydrophobicities of the TBS and pinacol protecting groups, the reaction was found to proceed very slowly with a very low yield of the desired product. A suspension formed during the reaction, performed in aqueous ACN, and the TBS protected starting material precipitated and therefore was present as a suspension. The reaction conditions, including the amount, ratio, or type of organic solvent (ACN, THF, DME), palladium ligands (TPPTS, TXPTS), and reaction temperature, were varied in an attempt to cope with solubility problems and improve reactivity and hence the overall yield but all modifications explored were found to be ineffective.

Substitution of 4-(TBS-O-methyl)-phenyl boronic acid (compound **(3)**) in this reaction scheme (Scheme 6.1) combined with increasing the ACN content in the coupling

was found to alleviate the solubility problem and lead to acceptable yields of the desired product (Table 6.1). Without pinacol protection, the hydrophobicity had been reduced. In addition, by increasing the organic content of the solvent (using 1:1 instead of 2:1 water:ACN), the solubility problem was completely eliminated.



Scheme 6.1 Synthesis of 4-(TBS-O-methyl)-phenylboronate pinacol ester

Other than the problem with solubility described above, the aqueous Suzuki coupling between 8-BrdG and aryl boronic acids has been proven to be an effective method to make C8-aryl modified dG. While the approach only provides medium to moderately high product yields (Table 6.1), it is a very clean and simple synthetic procedure and obviates the need for any intermediate protection or deprotection steps. Workup only involves precipitating the coupling products by the addition of 10% HCl to adjust the pH to between 6-7. This step is not always necessary as compound **6** precipitated as soon as the reaction mixture was diluted with water. In this particular case we found that by washing the precipitate with ethyl acetate was all that was required to remove excess starting material, compound **3**. Generally, it was found that precipitation

of the coupling products, without further purification, provided sufficiently pure material as confirmed by NMR and ESI-MS analysis of the coupling products.

Table 6.1 Synthesis yields of C8-aryl modified dG and derivatives prepared by Suzuki coupling.

p-Substituent on Aryl Adduct	Percentage Yield (%)		
	Suzuki Coupling	N2 Protection	5'OH Protection
COOH	82.6	95.7	30.9
CH_2OCH_3	67.1	87.6	54.5
CH_2OTBS	71.8	77.3	41.4

The protection reaction of the N2 amine of 2′-deoxyguanosine residue of the modified base produced good yields of the desired product (typically > 75%, Table 6.1). Based on ^1H NMR analysis, the nucleoside precursor was completely consumed resulting in only the production of N2 protected nucleoside and the by-product, DMF. We found that by precipitation of N2 protection product by the addition of water was an effective way to remove the DMF and thereby further purify the crude product, but this procedure resulted in a loss of product in the case of compounds **(8)** and **(9)** likely because these products were slightly water soluble. The product of *p*-carboxy derivative (compound **(7)**) was quite water soluble and thus could not be precipitated. Since DMF does not interfere with 5′OH tritylation, we did not typically remove it from a crude product prior to conducting this step.

Unlike the previous two reactions, the 5′OH protection reaction required column chromatography to isolate the trityl protected product. The yield was relatively low (30-50 %, Table 6.1) probably because of several factors including, product degradation, side reactions, and loss of sample during column chromatography. The glycosidic bond of C8-

aryl modified dG is weaken relative to the unmodified nucleoside, and is very acid sensitive. In addition, the DMTr protected 5′OH is, itself, an acid labile functional group which is easily cleaved off, reverting back to the unprotected 5′OH in the presence of protic acid. Based on TLC obtained during the reaction, the tritylation did not go to completion after 12 hours. Using extended reaction times, longer than 4 hours, resulted in the production of the undesired double tritylation product (tritylation of both 5′OH and 3′OH). Thus, in order to minimize by-product formation at the cost of product yield the reaction was stopped after 4 hours. Finally, as with the previously synthesized C8-arylguanosine derivatives, using acidic silica gel as a stationary phase causes degradation of the trityl protected product. Therefore, neutral or basic Al_2O_3 gel has to be used to minimize the degradation.

Among the three nucleosides prepared in this work, the yield of the carboxyphenyl derivative (compound **(10)**) was the lowest yield (Table 6.1). This is due to the purification process we had to use - precipitation instead of column chromatography. Because the carboxylate substituent strongly bound to the Al_2O_3, column chromatography purification was not possible. However, the carboxylate derivative **10** displayed poor solubility in dichloromethane (the reaction solvent) was low enough such that we were able to isolate it from crude product by precipitation. Although this method's effectiveness has appeared to be moderately low, it was more practical than column chromatography.

In the case of the *p*-hydroxymethyl derivative, the instability of TBS protecting group toward basic conditions may explain the low yield of **(12)**. The basicity of pyridine and TEA used in the reaction, to neutralize the HCl formed, may also have caused

deprotection of TBS group in **(12)** or the starting material. As expected the *p*-methoxymethylphenyl derivative did not have any significant solubility or stability issue, therefore, the product yield of **(11)** was, relatively higher.

The intermediates, especially after introduction of the aryl group, were prone to decomposition via cleavage of the glycosidic bond. The intermediates leading up to the phosphoramidite could be purified but, even with chromatography on basic alumina, significant decomposition occurred. Purification of phosphoramidite resulted in degradation of product and has shown to be unnecessary since crude phosphoramidites were effectively used to make the modified oligonucleotides. ^{1}H NMR analysis was used as a quick method to screen a crude phosphoramidite used in DNA synthesis. This is somewhat complicated by the fact that the chiral phosphoramidite group was introduced on the 3′OH and significantly increased the the multiplicity of peaks, especially sugar protons, as the number of diastereomers present doubled.

The phosphoramidites prepared as described were successfully used to make oligonucleotides but exhibited an unfortunate behavior of aggregate formation and precipitation in ACN, though this behavior was not seen with other solvents. Once formed, this material was insoluble in most aprotic solvents. The aggregate was not identified but we speculate that it may be related to the G-quartet. Guanosine and derivatives are known to form gels though gel formation tends to require the presence of a metal cation. While it is unknown what the constitution of the gel is, it is relevant to note here as it directly affects oligonucleotides synthesis.

Gel formation is relatively fast, occurring over a few minutes and this may hinder addition of the base for coupling during oligonucleotides synthesis by limiting the access

of phosphoramidite solution to CPG surface where base coupling takes place. It is unlikely that the manual addition procedure was directly responsible for the relatively low yields obtained as manual addition of unmodified bases did not decrease yields. Instead, the lower yield with the C8-arylmodified guanines was more likely due to the acid sensitivity of the glycosidic bond. Each base addition cycle on an automated DNA synthesizer exposes, albeit it briefly, the growing oligonucleotide to strong acid (TCA) which may cause partial decomposition.

Because of the possibility of incomplete extension and decomposition of oligonucleotides containing the C8-arylguanine modification during synthesis, purification was necessary. FPLC chromatograms of the crude CG^{8CPh} or CG^{8MMPh} oligonucleotides shows a series of minor peaks associated with incomplete sequences and degraded products that appear before the major peak of the desired product. In the case of CG^{8HMPh}, two large peaks were observed in the chromatogram and we believe them to be the peaks of TBS protected and unprotected CG^{8HMPh} oligonucleotides, which was subsequently proved to be true by ESI-MS analysis. The TBS protected CG^{8HMPh} was collected and re-reacted with NH_4OH to give CG^{8HMPh}.

As expected, the retention time of CG^{8MMPh} was shorter than CG^{8HMPh} because the methoxy residue decreases the hydrophilicity and thus the affinity for anion exchange resin relative to a hydroxyl group. CG^{8CPh} was expected to elute more slowly than the other two oligonucleotides due to the polar nature of the –COOH group and that, under the FPLC conditions, it is negatively charged (the FPLC conditions used to separate oligonucleotides based on charge and hydrophobicity). Interestingly, the results from the FPLC have shown otherwise as it was found that CG^{8CPh} eluted faster than CG^{8HMPh}. It is

possible that the conformation of the carboxy bearing oligonucletide is different that that with the hydroxymethyl derivative, perhaps screening it from the stationary phase.

Overall, in this study, we have demonstrated that our previously used methods can be considered a standard approach for the preparation of C8-aryl modified guanine phosphoramidites. Further, these phosphoramidites can be used for the preparation of oligonucleotides though the utilization of them requires manual addition. The possibility that they are sensitive to acid should be noted as it may impact on decisions as to where to place modified oligonucleotides in a sequence. Since automated DNA synthesis is typically from the 3′ end, decomposition will be minimized if the modified base is nearer the 5′ end, given that the option exists. According to FPLC and ESI-MS, the synthesized oligonucleotides were pure enough to use in structural and conformational analysis by CD and NMR.

6.2 CD Analysis and Molar Fraction Calculation

The global conformation of oligonucleotides can be determined from the shape of the CD curves. Therefore, the transition of B form to Z form of CG decamers can be observed as can the B/Z equilibrium constant using CD spectroscopy. The C8-arylguanine adducts have been shown to effect the B-Z equilibrium, shifting it toward the Z-DNA form. This can be demonstrated since the modified CG decamers required much lower salt concentrations as compared to the unmodified CG decamer. With the exception of C8-carboxyphenyl adduct, which is obviously the best at shifting the B-Z equilibrium toward the Z-DNA form of the CG decamer, the remaining two aryl adduct containing oligonucleotides were similar to the phenyl derivative. At the same time and

as will be discussed below, the NMR data and modeling results suggested that the three oligonucleotides studied here adopted conformations different from that adopted by the phenyl derivative. Here, we sought to put the effect on a quantitative basis and determine equilibrium constants and therefore thermodynamic parameters.

The molar fractions of three DNA conformations (random coil, B, and Z forms) that may be present in sample solutions of our modified oligonucleotides have been calculated from CD data. The effect of adducts on B-Z conversion has been determined with comparisons made based on the salt concentrations that are required to obtain CG decamers in 50 % in the Z form ($f_Z = 0.5$, Figure 4.13). The lower the salt concentration needed to achieve this indicates a greater effect on B-Z equilibrium. Based on the response to salt concentrations, then, CG^{8CPh} was found to be the most prone to form the Z-DNA conformation followed by CG^{8Ph}, CG^{8MMPh}, CG^{8HMPh}, and finally CG^{8Tol}.

The effect is partially due to destabilization of B-DNA form, as suggested from Figure 4.14, which illustrates the modified DNA have higher concentrations of the single-stranded DNA (ssDNA) forms under a given set of conditions (i.e., denaturation of the B DNA is occurring) than the unmodified DNA, especially at lower salt concentrations, conditions under which the B form should be predominant. To the extent that the ssDNA and Z forms are in equilibrium and that increasing the concentration of the ssDNA form will tend to drive the equilibrium toward the Z form, this may be one of the possible factors that affect the B-Z equilibrium. The position of the B-Z equilibrium is one measure of the effect of the aryl modification. However, it is a relative one and does not directly address if the effect is one of destabilizing the B-DNA form, stabilizing the Z-DNA form, or both. To explore this aspect, we used the melting temperature (Tm) data

(Figure 6.1) and, in particular, selected the conditions under which molar fraction was f_{ss} = 0.5. For the unmodified CG decamer, Tm continuously dropped as salt concentration increased. Ignoring the effect of salt concentration on Tm (for a B DNA, increasing the salt concentration tends to increase the Tm), the overall decline in Tm (e.g., above 500 mM NaCl) reflects the shift in the B-Z equilibrium toward the Z form, which predominates at high salt concentration. Thus, based on Tm, the Z-DNA form is less stable than the B-DNA form.

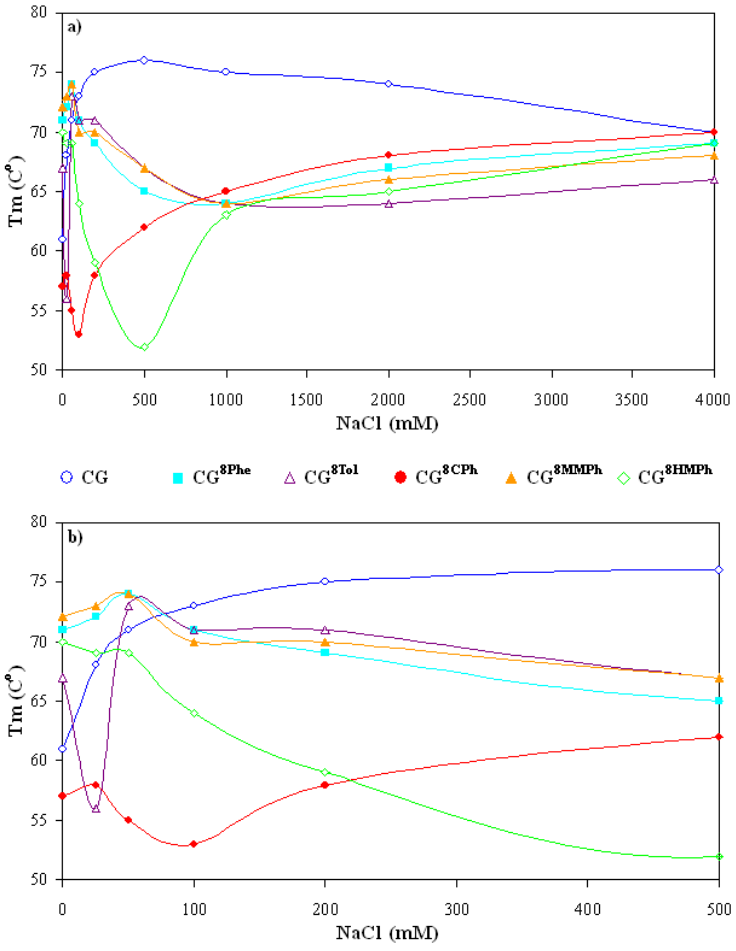


Figure 6.1 The Tm of the unmodified and modified CG decamers in phosphate buffer pH 7.4 with NaCl concentration of a) 0-4000 mM. The expansion of 0-500 mM is shown in plot b).

Introduction of the aryl adduct clearly altered the stabilities of B and Z-DNA. The effect was the most pronounced for the B form and the aryl group tended to destabilize as can be seen by considering the decline of Tm at low salt concentration for all modified CG decamers. The source of this destabilization, in part, is likely due to a steric effect of C8-aryl as there is limited space around C8-position of an *anti* G6 in B form of CG decamer. In this regard, the Tm of CG^{8CPh}, CG^{8Tol} and CG^{8HMPh} are unique and show a sharp and substantial decline, initially, as the salt concentration is increased and this is not observed for the CG^{8Ph} or CG^{8MMPh} oligonucleotides. This behavior does not seem to be consistent with being caused by only steric interactions to explain the Tm results since *p*-CH_3, *p*-COO^- and *p*-CH_2OH are all of similar size to the *p*-CH_2OCH_3 group. An electrostatic interaction may play an additional role to destabilize B form of CG^{8CPh} as, at pH 7.4, the carboxylic group will be negatively charged and may unfavorably interact with the phosphate backbone. While not charged, and potentially H-bonding to the phosphate backbone, the *p*-CH_2OH also bears lone-pairs and, if H-bonding is occurring, will also be more negatively charged than an isolated alcohol oxygen. The suggestion of electronic destabilization is supported by the molecular modeling results. Based on the most representative structures of B-CG^{8CPh} and B-CG^{8HMPh}, the nearest distance between negatively charged oxygens of *p*-COO^- and C5 phosphate oxygen was 3.4 Å while the neutral oxygen of *p*-CH_2OH was even closer to C5 phosphate oxygen with the distance of 2.8 Å (Figure 6.2). The distances were far more than 10 Å in Z form of both oligonucleotides, diminishing the electronic effect that may be existed in B form. Note that the lack of negatively charge of *p*-CH_2OCH_3 and longer distance from C5 phosphate may explain why the Tm of CG^{8MMPh} does not follow this trend (though it may

and additional points near 0 mM NaCl need to be collected). The most representative structure and helical parameters obtained from structure analysis in computational study does support the idea of B-DNA destabilization. However, electrostatic interactions like those described for B-CG^{8CPh} and B-CG^{8HMPh} clearly are not possible for CG^{8Tol} and other or additional factors must be involved.

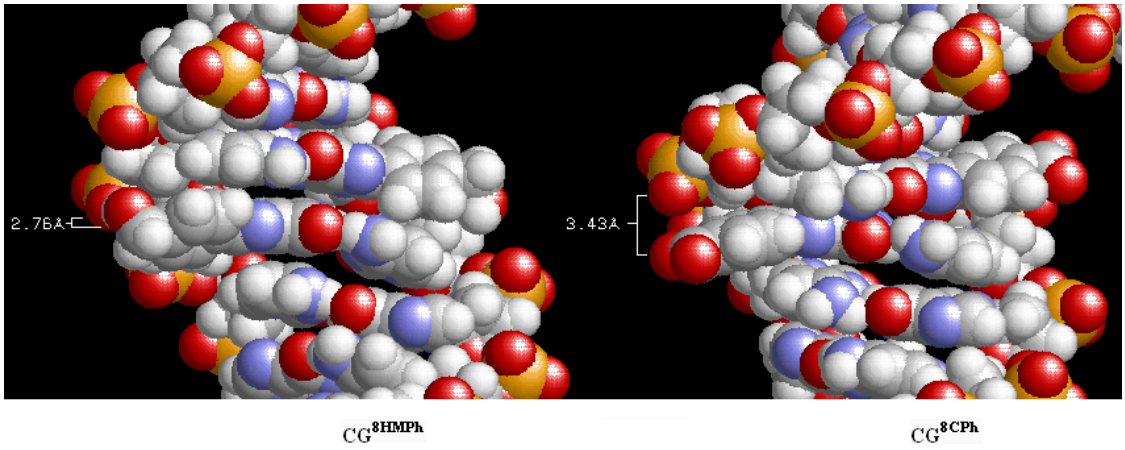


Figure 6.2 Snapshot shows possible electrostatic interaction between C5 phosphate and the *p*-substituents of aryl adducts on CG^{8HMPh} (left) and CG^{8CPh} (right).

Although the destabilization of B form can be extrapolated from the Tm data, the effect of the aryl adduct on Z-DNA stability is more difficult to ascertain. This is because as the temperature increases and the Tm approached, an ever increasing amount of the Z DNA form converts to the B-DNA form prior to melting. Therefore, it is not necessarily correct to compare the stability of Z-DNA based on the observed Tm only, even at higher salt concentrations. Nevertheless, if it can be assumed that at 4000 mM NaCl, that most of the modified oligonucleotides are in the Z-DNA conformation during melting, then some comparisons can be made. In particular, if there is destabilization of the Z-form then the order is CG^{8Tol} > CG^{8MMPh} > CG^{8HMPh} > CG^{8Ph} > CG^{8CPh}.

6.3 NMR Analysis

In addition to the global conformation of DNA that can be detected by CD, the determination of local conformation can be obtained from NMR analysis. The results from the NMR studies were mainly in agreement with the CD data. Thus, they confirm that at low salt the modified CG decamers were in B form while Z form was a predominant conformation in samples containing high salt concentrations. The latter is demonstrated by the strong nOe cross peak of H8/H-1′ suggesting that the purines (dG) were adopting the *syn*-conformation as they are in Z-DNA.

The chemical shifts of C8-aryl adduct of G6 (Table 6.2) have implications for what is the local structure of G6 in the B and Z forms. The aryl adduct protons of the B form were found to be upfield relative to these protons in the Z form. This indicates that, in the B conformation, G6 may be positioned in such a way that the aryl ring of the adducts are under an influence of the π system of the 5′ pyrimidine ring (C5). This would produce a shielding effect on aryl adduct protons, shifting them upfield. This suggestion is supported by the molecular modeling results that show partial stacking/overlapping aryl ring of the C8-aryl adduct located at the G6 and the pyrimidine at C5 in B form of modified CG decamers. Unlike B forms, the aryl adducts in Z forms are pointed away from and are entirely outside of the helical axis and therefore cannot be shielded by neighboring bases.

It is also key to note that there were five distinct chemical shifts observed for the phenyl adduct in B form. This is not what was observed for the new adducts, rather the protons located at both the *ortho* and *meta* protons relative to the *p*-substituent were equivalent and displayed a the typical AA′XX′ pattern of a 1,4-disubstituted benzene ring

and not four distinct resonances. One explanation for this is that there is free rotation about the aryl-dG bond in the *p*-substituted aryl systems such that the two *ortho* or *meta* protons can interconvert, but not in the phenyl substituted system. Based on structural analysis in computational study, the helical parameter suggested more stacking between phenyl adduct and pyrimidine of C5 which may constraint the phenyl adduct so that it can not rotate as fast as other adducts.

Table 6.2 ^{1}H NMR assignment of C8-aryl G6 adduct in CG decamers.

DNA	Chemical shifts of aryl adducts protons (ppm)	
	B form	Z form
CG^{8Ph}	6.35 ,7.04 ,7.28 ,7.32 ,7.33	7.55, 7.65, 7.84
CG^{8Tol}	6.28, 6.86	7.64, 7.69
CG^{8CPh}	6.72, 7.79	7.81, 8.16
CG^{8MMPh}	6.38, 7.08	7.74, 7.77
CG^{8HMPh}	6.37, 7.09	7.73, 7.76

The assignment of non-exchangeable protons to each base in CG decamer sequences has shown some interesting trends in chemical shifts that contain structural implications. The aryl adducts have structurally affected their neighboring base as we observed from the altered chemical shifts. In case of the unmodified B-CG (Figure 6.3a), the chemical shifts for most of a given type of proton were fairly constant except at the ends were fraying effects modulate shifts, and are consistent with well oriented base pairs in a double helix. In contrast, the C8-aryl adduct on G6 of, for example, B-CG^{8MMPh} (Figure 6.3b), caused discontinuities in chemical shift of all protons of G6 and C5 (G15 and C16) suggestive of a local disruption in structure. This also supports the CD results that show the aryl adduct destabilized B form.

Molecular modeling studies suggested the same idea and have been used to help explain some of the abnormality in chemical shifts of the modified CG decamers. For instance, in the model of the B-CG^{8MMPh}, C5:H6 and C5:H-2′ are located right above the aryl adduct on G6 which should position them such that they should be shielded and be upfield shifted as observed. Likewise, the unusual downfield H-1′ of C7 could be a consequence of a remarkably close distance (2 Å) between the H-1′ and G8:N7, which were generally farther (5 Å) for other cytosines. All aryl adducts (except for *p*-carboxyphenyl due to unavailable NMR assignment of B-CG^{8CPh}) have been show to cause similar patterns in chemical shift plots of protons in oligonucleotides in B form.

The conformational transition of B to Z form requires rearrangement of the DNA structure that is reflected in the change in proton chemical shift. By comparison of proton chemical shift plots of the B-CG (Figure 6.3a) and Z-CG^{8CPh} (Figure 6.4a) or Z-CG^{8MMPh} (Figure 6.4b), we can observe the differences in chemical shift between B and Z-DNA. The most notable ones were including the downfield shifts of G:H-1′, C:H5, and the upfield shift of C:H-2′. The key to these chemical shift changes is based on the transition of *anti*-dG to *syn*-dG as occurs when B-DNA converts to the Z conformation. In B form, dG bases are in the *anti*-conformation such that H-1′ of dG are positioned deep inside the helix and are probably shielded by neighboring pyrimidines. Similarly, H5 of dC experience a shielding effect caused by the purine of 5′ dG. Once the conformation of dG reverses to *syn*, the shielding effects are no longer possible resulting in downfield chemical shifts of G:H-1′ and C:H5. On the other hand, the *syn* conformation of dG places H-2′ of dC closer to N2 amino group of 5′ dG causing C:H-2′ to be slightly upfield.

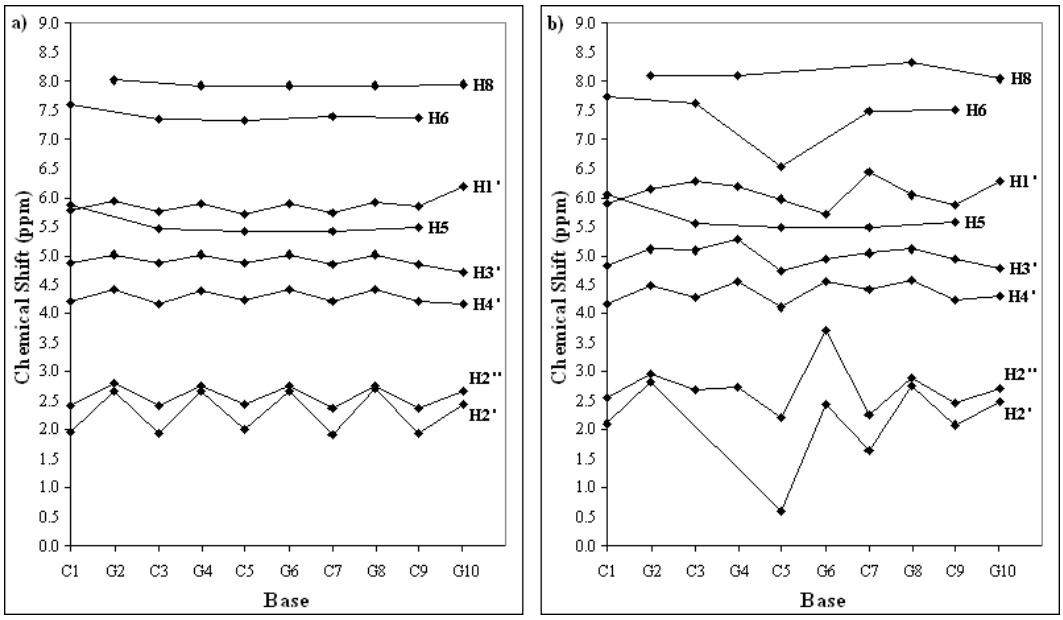


Figure 6.3 The plots of non-exchangeable proton chemical shifts of each base in a) B-CG and b) B-CG^{8MMPh}

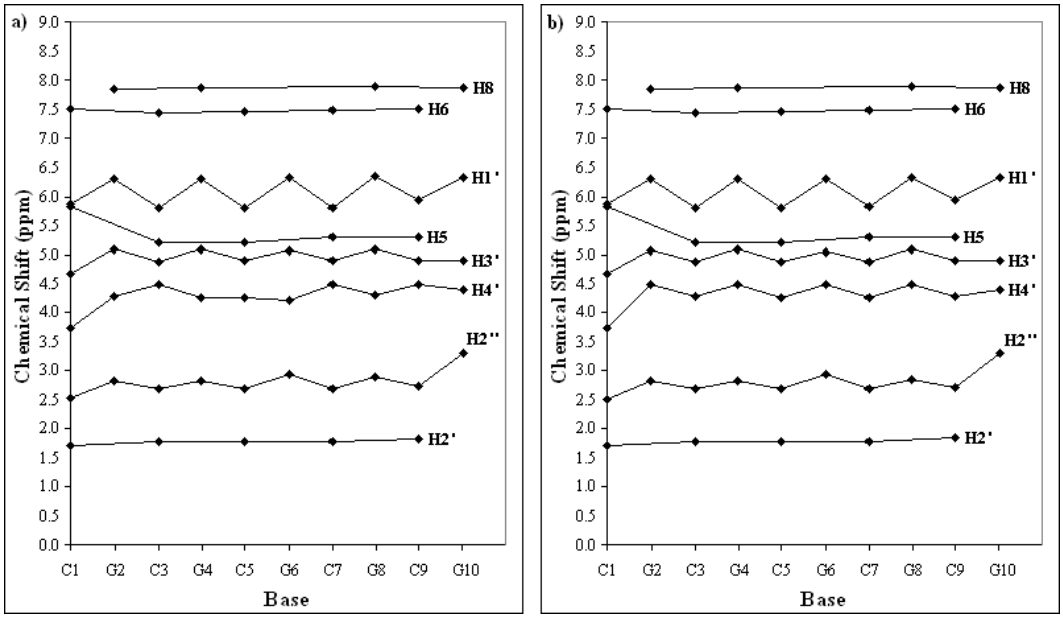


Figure 6.4 The plots of non-exchangeable proton chemical shifts of each base in a) Z-CG^{8CPh} and b) Z-CG^{8MMPh}

As we can see from the continuity of plot of chemical shift in Z form, the aryl adducts do not seem to disrupt the structure of Z form as they did in B form. This remark is generally supported by the model of modified CG in the Z-DNA conformation with the exception of Z-CG^{8CPh}. In this case, the aryl adduct is pointed away from the helical axis due to *syn*-glycosidic bond of G6 that position such that the steric interaction between H8 and H-2′/2″ was reduced. Consequently, this would prevent the aryl adduct from interacting and interfering with base pairing, unlike what occurs when this adduct adopts the B-DNA form. NMR analysis has suggested additional structural information that the aryl adducts favor Z over B conformation.

6.4 Computational study

The most representative structures and helical parameters have shown that the aryl adducts, in general, distorted the B conformation. Based on general considerations, the bulkiness of aryl adducts is the likely cause of the structural deformations. However, the simulations of various aryl modified CG decamers have pointed toward the possibility that other interactions may also play important roles with respect to the effect of them on the B-DNA conformation, both locally and globally. A major factor is the occurrences of stacking interactions between the C8-aryl adduct on G6 and pyrimidine ring of C5 was found to be energetically possible, especially in the case of the phenyl adduct. Structurally, this interaction requires, locally, increase rise and negative slide of the base pair which, in turn, stretches the duplex. These changes decrease the overall contribution of base-base stacking and destabilize the B-DNA conformation.

There may also be electrostatic repulsion between the *p*-substituent that further decreases the overall stability of the B-DNA. All of the aryl adducts have been shown to cause compression of the minor groove (decreased width). Decreasing the minor groove width causes the phosphate groups to be closer to one another causing an unfavorable electrostatic interaction. In addition, depending upon the specific adduct, additional unfavorable electrostatic factors may further destabilize the B-DNA form as in the case of the *p*-COO^- derivative in which the negatively charged carboxylate group is position near the negatively charged phosphate backbone. This additional factor is the likely cause of the low stability of B-CG^{8CPh} that observed in Tm profile (Figure 6.1).

The simulations of the C8-arylguanine adducts in the Z-DNA conformation have not elucidated the order of stability experimentally determined. Except for CG^{8CPh} and CG^{8MMPh}, in which the aryl adducts seemed to have caused some effects on duplex structure. Nevertheless, with current data from the structure analysis, the computational studies have not fully explained how the aryl adducts are effecting Z-DNA stability.

The free energy calculation was conducted in an attempt to gain more information regarding Z-DNA stability, but the free energy calculations based on MM_PBSA or sietraj are not in agreement with the experimental results. There are several potential reasons for this that were previously discussed and additional reasons, such as the effect of the choice of sampling (i.e., which structures are used to calculate the free energy) which has been shown to provide different free energy results. Also, it is possible that the length of simulations were insufficient. This is particularly true for the Z-CG^{8MMPh} as the RMSd plot indicated that the structure was still drifting from the starting structure and changing helicoidal parameters determined with Dials & Windows, even after 4 ns of

MD. Revision or continuation of the simulation(s) of the Z-DNAs may be required to sort out the problem with the computation and thereby resolve the differences observed between the experimental and computational results.

6.5 Thermodynamic study

The B-DNA destabilization effect of the aryl adducts seems to be clear based on the CD, NMR, and computational studies. The effect on the Z-DNA form is still ambiguous. Thermodynamic studies may be able to aid in understanding what is driving the shift in the B-Z equilibrium as a function of the aryl adducts. The van't Hoff plot of the B-Z equilibrium has been made to obtain thermodynamic parameters including ΔH, ΔS, and ΔG[20,21,85]. The system equilibrium has been simplified to ease the calculation as depicted in Figure 6.5, and assumes that B, Z, and ssDNA are the only species present and simplifies the analysis by assuming that only the B and Z forms interconvert ($K_{B\text{-}Z}$). The alternative is that this is a three species equilibrium (i.e., B/ss followed by ss/Z rather than B/Z) but is much more difficult to fit.

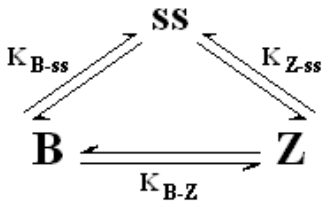


Figure 6.5 A diagram shows equilibrium between B-, Z-, and ssDNA

The van't Hoff equation was derived as following

$$\Delta G = \Delta H - T\Delta S \qquad \text{Equation 6.1}$$

$$\Delta G = -RT\ln K_{B\text{-}Z} \qquad \text{Equation 6.2}$$

$$\ln K_{B\text{-}Z} = -\Delta H/RT + \Delta S/R \qquad \text{Equation 6.3}$$

$$\ln(f_Z/f_B) = -\Delta H/RT + \Delta S/R \qquad \text{Equation 6.4}$$

$$\ln(f_B/f_Z) = \Delta H/RT - \Delta S/R \qquad \text{Equation 6.5}$$

Based on Equation 6.5, by plotting $\ln(f_B/f_Z)$ against 1/T, ΔH can be obtained from slope (ΔH/R) and ΔS can be obtained from intercept (-ΔS/R), where R is the gas constant (1.9872 cal $mol^{-1}K^{-1}$). The van't Hoff plots of B-Z transition of the unmodified and modified CG decamers are shown in Figure 6.6 and the thermodynamic parameter are reported in Table 6.3. From the thermodynamic data, the B-Z conversion was mainly enthalpy driven. The aryl adducts make the transition to the Z-form more favorably by lowering energy of the Z-DNA form based on significant enthalpy change, even though the B-Z conversion process of the modified oligonucleotides reduces the entropy of the system. Although the negative ΔG of B-Z conversion indicates that the process was energetically favorable (which is obvious), it does not elucidate the source of the stabilization.

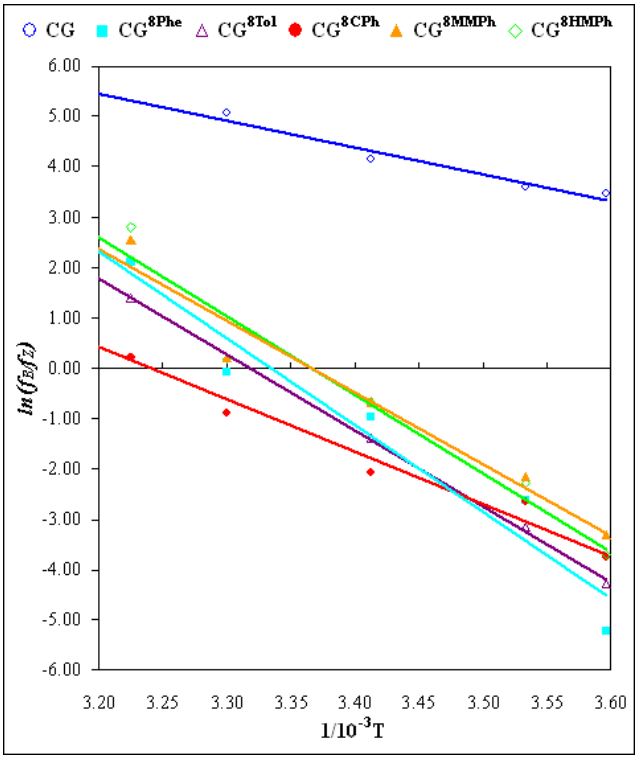


Figure 6.6 The van't Hoff plot of B-Z transition of the unmodified and modified CG decamer in phosphate buffer pH 7.4 with 500 mM NaCl.

Table 6.3 Thermodynamic parameters for B-Z transition of the unmodified and modified CG decamer in phosphate buffer pH 7.4 with 500 mM NaCl at 298°K (25°C).

DNA	Slope	Intercept	R^2	ΔH (cal mol^{-1})	ΔS (cal mol^{-1} $K^{\circ -1}$)	ΔG (kcal mol^{-1})
CG	-5.33	22.52	0.95	-10594	-44.75	2.74
CG^{8Ph}	-16.30	52.65	0.99	-32389	-104.62	-1.21
CG^{8Tol}	-15.06	49.98	1.00	-29933	-99.33	-0.33
CG^{8CPh}	-10.85	33.04	0.99	-21553	-65.66	-1.99
CG^{8HMPh}	-14.63	47.74	0.99	-29071	-94.86	-0.80
CG^{8MMPh}	-13.30	43.39	0.98	-26426	-86.23	-0.73

To further explore the idea that the aryl adducts stabilize the Z-DNA form and thus work synergistically with their effect of destabilizating the B-form, ΔG of Z and B-DNA duplex binding have been estimated using the same method as used for B-Z

conversion. A decline in ΔG (i.e., more negative) would be expected if the duplex in the Z-form was stabilized in the presence of an aryl adduct, since this means higher energy is required to denature double strand DNA to single strand DNA. On the other hand, increase in ΔG means CG decamer duplex was destabilized by an aryl adduct. The thermodynamic parameter of B-DNA and Z-DNA duplex binding are shown in Table 6.4 and 6.5.

Table 6.4 Thermodynamic parameters for B-DNA duplex binding (ss→B) of the unmodified and modified CG decamer in phosphate buffer pH 7.4 with 500 mM NaCl at 298°K (25°C).

DNA	Slope	Intercept	R^2	ΔH (cal mol^{-1})	ΔS (cal mol^{-1} $K^{\circ -1}$)	ΔG (kcal mol^{-1})
CG	-21.22	59.65	0.98	-42176	-118.53	-6.85
CG^{8Ph}	-7.11	20.65	0.95	-14123	-41.03	-1.90
CG^{8Tol}	-6.65	19.16	0.91	-13213	-38.07	-1.87
CG^{8CPh}	-7.28	21.78	0.93	-14475	-43.28	-1.58
CG^{8HMPh}	-7.88	23.35	0.96	-15655	-46.40	-1.83
CG^{8MMPh}	-8.45	24.44	0.96	-16793	-48.58	-2.32

Table 6.5 Thermodynamic parameters for Z-DNA duplex binding (ss→Z) of the unmodified and modified CG decamer in phosphate buffer pH 7.4 with 1000[a] mM NaCl at 298°K (25°C).

DNA	Slope	Intercept	R^2	ΔH (cal mol^{-1})	ΔS (cal mol^{-1} $K^{\circ -1}$)	ΔG (kcal mol^{-1})
CG	-16.94	52.58	0.99	-33655	-104.48	-2.52
CG^{8Ph}	-27.86	85.28	0.97	-55353	-169.46	-4.85
CG^{8Tol}	-20.19	62.37	1.00	-40130	-123.94	-3.20
CG^{8CPh}	-15.46	46.06	0.95	-30718	-91.52	-3.44
CG^{8HMPh}	-27.55	84.17	0.97	-54743	-167.26	-4.90
CG^{8MMPh}	-19.04	57.74	0.99	-37840	-114.74	-3.65

[a] Data points at 1000 mM were used in van't Hoff plot of Z-DNA → ssDNA instead of at 500 mM due to the f_Z data of the unmodified CG that available only at higher salt concentration.

By comparing ΔG of the modified CG decamers to the unmodified one, we have again confirmed that the aryl adduct destabilized B form and by far B-CG^{8CPh} was the least stable one as expected from experimental and computational studies. More importantly, we have finally demonstrated that the aryl adducts actually stabilized Z-DNA as ΔG of Z-DNA binding were lower in case of the modified CG decamers.

The aryl adduct stabilizes the Z-DNA by reducing system enthalpy and suggests, from the thermodynamic parameters obtained and shown in Table 6.5, except in the case of the *p*-carboxyphenyl adduct. This unique behavior was actually observed in the most representative structure in which the Z-CG^{8CPh} seemed to bind relatively loose compared to the rest of CG decamers.

Note that Z-DNA is known to be unusually rigid[20,21] as compared to its counterpart B conformation or other DNA conformations. The decrease in entropy of the modified CG decamers compared with the unmodified one may suggest that the aryl adducts have affected the structure of Z-DNA such that it is even more rigid than the unmodified version. The computational study has shown us that the Z form of CG^{8CPh} seemed to be more flexible compared to the rest of modified CG decamers. This notion corresponds to the free energy estimation that formation of Z-CG^{8CPh} has the smallest entropic penalty. The unique electronic effect of *p*-COO$^-$ is likely involved in driving structural change in a way that lowered rigidity of Z form compared to other aryl adducts.

6.6 Relevance of the B-Z Transition in Aryl Hydrazine Carcinogenesis

To evaluate the role of B-Z transition in aryl hydrazine carcinogenesis, we have set out to make the modified CG decamers with selected C8-arylguanine adduct. By linking the effect of the aryl adduct on Z-DNA stabilization/B-DNA destabilization and carcinogenic profile of the aryl hydrazine that serves as the precursor to adduct formation; we hoped to show the relevance of B-Z equilibrium in carcinogenesis will be revealed. While it has been shown in our study that all studied carcinogenic aryl hydrazines can generate aryl adducts that stabilized Z-DNA and/or destabilized B-DNA and shift the B-Z equilibrium toward the Z-DNA form, we also hoped that a ranking of the effect would correlate with the carcinogenicity of the aryl hydrazines.

There have been numerous studies of the mutagenicity and/or carcinogenicity of aryl hydrazines and related chemicals. These studies have been conducted in a wide range of several cell lines and animals[4]. Although the pool of carcinogenic data is large, a consistent data set (e.g., same dose, species, etc.) in which the carcinogenicities of various aryl hydrazines were compared under identical conditions is unavailable. Usually different models or conditions were used for each study and a clear ranking is difficult to create. The most comprehensive carcinogenic profile was from Lawson, T., et al., work that was conducted in collaboration with our lab in 1995[46]. In this study, the mutagenicities of selected aryl hydrazine metabolites, including MBD, HMBD, MMBD, and CBD, were measured in TA102 and V79 cells. The result on TA102 (Figure 6.7) suggests that the mutagenic potency of arenediazonium ions was in order of MBD > HMBD > MMBD > CBD, which turned out completely opposite to the trend the resulting adducts derived from the benzenediazonium ions have on the B-Z equilibrium found in

this study. Note that MBD gives *p*-tolyl adduct while HMBD, MMBD, and CBD give *p*-hydroxymethylphenyl, *p*-methoxymethylphenyl, and *p*-carboxyphenyl adducts respectively.

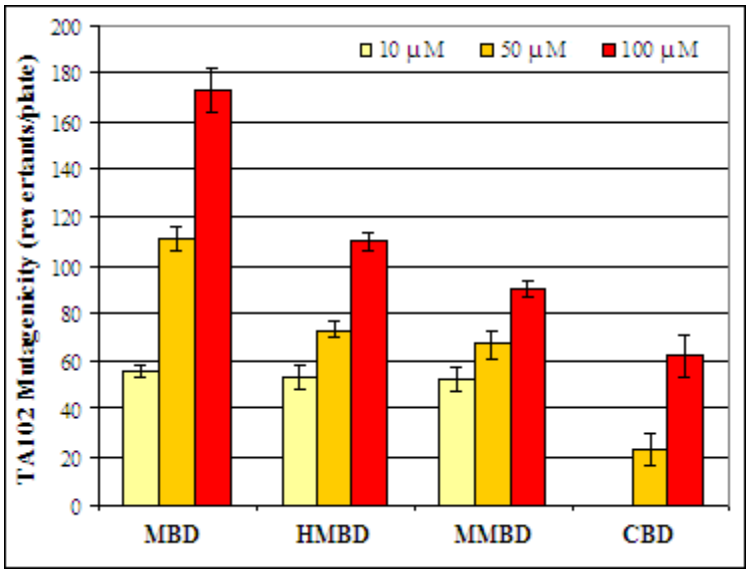


Figure 6.7 Mutagenicity of arenediazonium ions in TA102 cells

Considering that the known mechanism of mutagenesis in TA102 model was based on the frame shift caused by mutation specifically at a 'hot spot' comprised of five consecutive A-T base pairs[113], the data seemed to be less relevant and may not be directly applicable to our work that used CG decamer sequence as a testing model of B-Z equilibrium.

The result obtained from the V79 assay, a more difficult but more relevant assay (mammalian cells rather than bacteria). The assay itself is run in Chinese hamster ovarian cells and is based on the mutation of the hypoxanthine-guanine phosphoribosyl transferase (HGPRT)[114,115]. The mutagenicity of the arenediazonium ions based on this mutational analysis are shown in Figure 6.8 and show a different pattern of mutagenicity

in which MBD is now the weakest mutagen and the order of the remaining three arenediazonium ions is as observed in the Ame's assay.

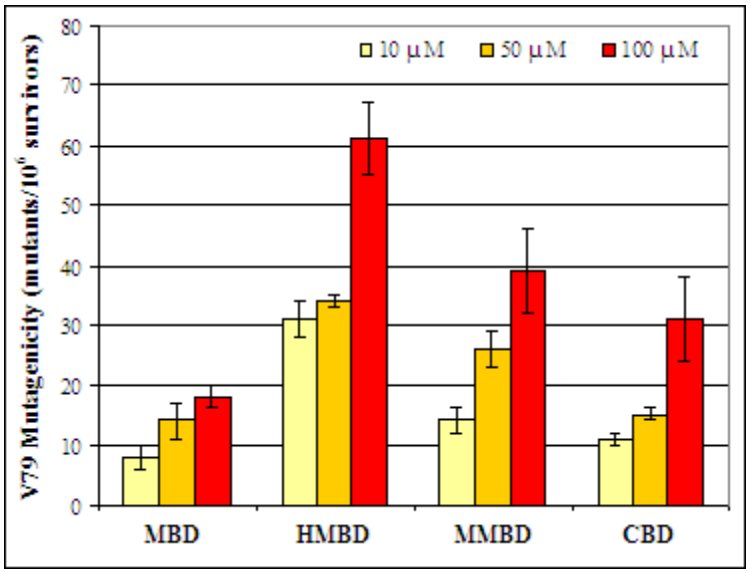


Figure 6.8 Mutagenicity of arenediazonium ions in V79 cells.

The mutagenic trend of arenediazonium ions in the V79 assay was better correlated with the results we had obtained regarding the Z-DNA stabilization effect observed for the modified oligonucleotides we had examined (*p*-hydroxymethylphenyl > *p*-methoxymethylphenyl > *p*-carboxyphenyl > *p*-tolyl) though the *p*-carboxyphenyl was out of order depending upon whether the effect of salt or Tm was used for the comparison. However, if the effect of the adducts on B-DNA stabilization is included, the mutagenic trend seemed to be unrelated to the effect of the adducts on B-Z equilibrium. The situation is further complicated if the stability of arenediazonium ions (tested in aqueous solution) is taken into account. If this factor is included then a possible correlation between mutagenicity and the B-Z equilibrium can be made. In particular,

arenediazonium ions are highly electrophilic and decompose very fast in cells culture media. The stability of arenediazonium in Williams medium E (WE) used in V79 study decrease in the order HMBD > MBD > MMBD > CBD (29%, 26%, 20%, and 14% remaining after 16 hrs in WE). The corresponding *p*-substituted phenols form as the arenediazonium ions decomposed. The degradation of the tested chemical in WE could greatly affect their mutagenicity and should be considered along with the reported data.

The adjusted trend (Figure 6.9) on mutagenicity of aryl hydrazines was generally aligned with the trend observed for the effect of the aryl adducts on the B-Z equilibrium. The tolyl adduct, which had the least effect on B-Z equilibrium forms from MBD, the weakest mutagen in V79 study. In addition, the V79 mutagenicity of HMBD was significantly higher than MMBD at a 10 μM dose, but the difference disappeared when tested at higher concentrations (50 μM and 100 μM dose) as the *p*-hydroxymethylphenyl and *p*-methoxymethylphenyl adduct have similar effect on B-Z conversion. Finally, as seen in Figure 6.8, the mutagenicity of CBD was relatively low, but after the stability of CBD in WE was considered, CBD was among the strongest mutagenic in the series and correlates with the result that *p*-carboxyphenyl adduct was the best at promoting B-Z transition.

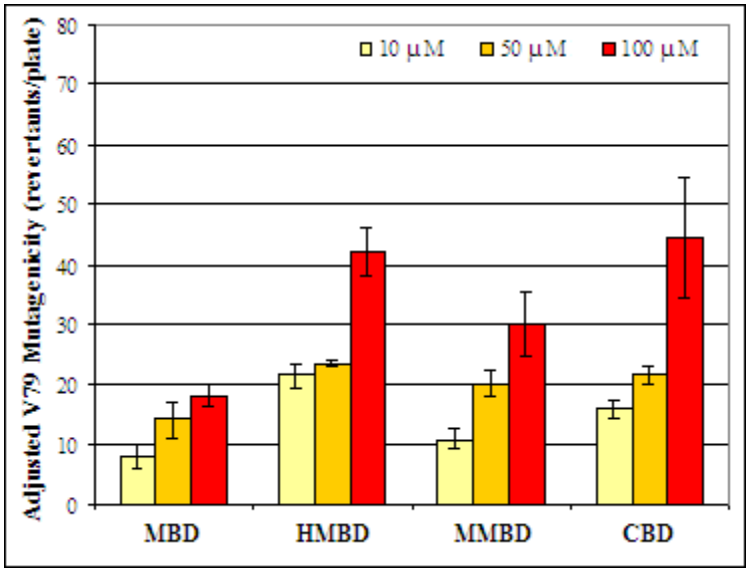


Figure 6.9 Mutagenicity in V79 cells factoring with arenediazonium ions stabilities.

An additional factor that needs to be considered is repair. While this is not an issue in the Ame's assay, in other settings, the rate of repair of damaged DNA (e.g., following adduct formation) is of considerable importance. A famous example of this is the O6-methyl and N7-methyl guanine adducts that form from N,N-dimethylnitrosamine. The latter is formed *in vivo* in amounts that are ten times that of the former[116]. However, the N7-methyl adduct is rapidly repaired while the O6-methyl adduct persists[117]. In the present case, the adducts all cause some degree of distortion of the DNA they are in. The tolyl adduct causes the greatest amount of distortion and is likely to be 'seen' by repair enzymes much more readily than the *p*-hydroxymethylphenyl adduct which produce only minor distortion of the DNA. Thus, how long the adducts persist may also play a significant role in the carcinogenicity of the parent aryl hydrazines.

Overall, it is unlikely that a simple correlation will be found between, for example, the Ames' Assay, V79 assay, repair assays, etc. Instead, all data, including the

conformational effects of the adducts will need to be considered to ascertain the role the B-Z equilibrium may play in aryl hydrazine mutagenesis and/or carcinogenesis. At the same time, our data do suggest the possibility of a correlation between aryl hydrazine carcinogenesis and B-Z equilibrium and therefore provides reasonable cause to continue with studies in a more biological setting (cellular level or *in vivo*) to more clearly determine the relationship between carcinogenicity and the effect of the aryl adducts on B-Z equilibrium.

6.7 Conclusion

Aryl hydrazines and related compounds have been known to be carcinogens several decades. Likewise, they have been extensively studied yet a mechanism or mechanisms of carcinogenesis is still unclear. The DNA aryl purine adducts generated from reactive metabolites of aryl hydrazines, including arenediazonium ions and aryl radicals, are suspected to be involved in or the cause of carcinogenesis like other adducts generate from non-hydrazine carcinogens (aminofluorene[17], PAHs[48], etc.). Several possible effects of the aryl adduct on DNA that may lead to genetic alterations have been suggested, among them the effect on the B-Z equilibrium as caused by aryl adducts has caught our interest.

The biological relevance of Z-DNA has been extensively studied and debated especially during the 1980's, though the debate continues and the scientific community as since largely ignored Z-DNA in a biological context. A similar history occurred for triplex DNA as when if was first discovered in 1957 it was consider an anomaly and nearly 30 years passed before a biological role was found. It was not until recently that Z-

DNA regained biological interest as Z-DNA formation has been shown to be an intermediate in activation process of CSF-1[37] which indicates possible biological role of Z-DNA in gene transcription. In addition, the mutagenic potential of Z-DNA has been proposed based on the result that links Z-DNA to DNA damage such as genetic deletions in mammalian cells[45]. These studies have pointed out the possibility that unregulated Z-DNA formation could consequently trigger genetic alteration and lead to carcinogenesis. With the possibility that aryl adducts formed from aryl hydrazines may promote B-Z transition like the well-known C8-bromo or C8-methyl guanine adducts, we have proposed that aryl hydrazine carcinogenesis may occur through altering the B-Z equilibrium facilitated, in the case of aryl hydrazines, by the aryl adduct. Therefore, we have set out to investigate 1) the effect of aryl adducts formed from aryl hydrazines on B-Z DNA equilibrium and 2) the relevancy between Z-DNA formation and aryl hydrazine carcinogenesis.

To this end, the CG decamer has been selected as the model system to investigate the effect of C8-arylguanine adducts on B-Z transition. The unmodified and modified CG decamers, including CG^{8Ph}, CG^{8Tol}, CG^{8CPh}, CG^{8MMPh}, and CG^{8HMPh}, were successfully made through Suzuki coupling and phosphoramidite chemistry. Structural and conformational analyses of the synthesized oligonucleotides, conducted by using CD and NMR, indicates that C8-arylguanine adducts have facilitates Z-DNA formation. The quantitated effect of adducts on B-Z transition estimated from salt concentration required to have CG decamers in Z form at a level of 50% suggests the preference on Z-DNA formation is in the order $CG^{8CPh} > CG^{8Ph} > CG^{8MMPh} > CG^{8HMPh} > CG^{8Tol}$. The melting temperature data shows that the adducts destabilize the B-DNA form. The

thermodynamic studies have confirmed the B-DNA destabilizing effect and show a simultaneous stabilizing effect on the Z-DNA form is caused by the C8-arylguanine adducts.

Computational studies have been used to understand and explain the local and global effects of the aryl adducts that drive B-Z conversion. The steric effect from C8-arylguanine adduct related to the *anti* dG form and the stacking interaction observed between the aryl adduct and C5 pyrimidine are most likely the main cause of the observed distortion seen for the B-DNA form and indicate a destabilizing effect on this conformation. Depending upon the adduct, there may be addition destabilizing effects. For example, the carboxylate group of the CG^{8CPh} oligonucleotide is also destabilized by electrostatic repulsion between the *p*-COO^- and the phosphate backbone and is likely the explanation for its unique effects relative to the other adducts on B-DNA destabilization. With current data, the Z-DNA stabilization effect of the aryl adducts can not be clearly explained except that in general, the adducts do not seem to interfere with the Z-DNA structure and therefore are not expected to decrease the stability of Z form like they do in B form.

We expected to find a correlation between carcinogenic potential of an aryl hydrazine and the effect on B-Z equilibrium. However, given the multitude of other factors that may impinge on the carcinogenicity of aryl hydrazines discussed above and additional factors such as absortion, distribution, etc., a perfect correlation is unreasonable to expect. However, based on the consideration noted above, CBD can be argued to be the most potent carcinogen followed by HMBD, MMBD, and MBD and this trend generally correlated with the trend observed for the position of the B-Z equilibrium

of CG^{8CPh}, CG^{8HMPh}, CG^{8MMPh}, and CG^{8Tol}. Based on the knowledge that all C8-arylguanine adducts formed from carcinogenic aryl hydrazines and that these adducts drive equilibrium toward the Z conformation, a better than expected correlation with mutagenicity and perhaps carcinogenicity has been revealed. Thus, this study has shown a potential connection between Z-DNA formation and aryl hydrazine carcinogenesis opening the possibility to advance the investigation to the next level.

6.8 Future Direction

The current data from this study suggests a possible link between Z-DNA and carcinogenesis by aryl hydrazines. Additional studies, however, will be required in order to obtain sufficient data to strongly support the proposed hypothesis.

In vitro studies have shown that the aryl adducts affect the relative stability of B and Z-DNA forms such that they shift the equilibrium toward the Z-DNA form. At the same time, we can not be certain that the aryl adducts will have the same effect *in vivo*. Although it may be true that some factors in the physiological environment such as Z-DNA binding proteins and other cations could promote Z-DNA formation, several factors may limit the formation of Z-DNA in physiological condition such as relatively low salt concentration, abundance of Z-DNA forming sequences, etc. Therefore, investigation of Z-DNA formation *in vivo*, as a result of adduct formation is needed. The presence of Z-DNA in physiological system is generally transient which will make observing it *in vivo* more difficult. No matter, it has been shown to be detectable by several methods including Z-DNA specific antibody binding assay[35,118-120], reaction with DEPC[37,121], and enzyme restriction of the B-Z junction[45]. In addition, detection of the binding to Z-DNA

specific proteins such as ADAR1[42,98], DML1[98,122], or E3L[123-125] would be a probable alternative or at least a first step toward detection of Z-DNA *in vivo*.

Ultimately, a more relevant, comprehensive carcinogenic profile of aryl hydrazines or arenediazonium ions from *in vivo* studies that incorporates the detection of Z-DNA formation as an intermediate process will be necessary to establish the correlation between Z-DNA and aryl hydrazines carcinogenesis.

REFERENCES

1. Toth, B. A review of the natural occurrence, synthetic production and use of carcinogenic hydrazines and related chemicals. *In Vivo* **2000,** *14* (2), 299-319.

2. Toth, B.; Gannett, P. Agaricus bisporus: an assessment of its carcinogenic potency. *Mycopathologia* **1993,** *124* (2), 73-77.

3. Walton, K.; Coombs, M. M.; Walker, R.; Ioannides, C. Bioactivation of mushroom hydrazines to mutagenic products by mammalian and fungal enzymes. *Mutat. Res.* **1997,** *381* (1), 131-139.

4. Toth, B. *Hydrazines and Cancer: A Guidebook on the Carcinogenic Activities of Hydrazines, Related Chemicals, and Hydrazines-Containing Natural Products;* Overseas Publisher Association: Amsterdam, The Netherlands, 2000.

5. Lawson, T.; Chauhan, Y. Metabolism of arylhydrazines by mouse liver mixed-function oxidase in vitro. *J. Agric. Food Chem.* **1985,** *33*, 218-219.

6. Lawson, T. Metabolism of arylhydrazines by cytochrome P-450 mixed function oxidases and prostaglandin(H)synthase from mouse lungs. *Cancer Lett.* **1987,** *34* (2), 193-200.

7. Gannett, P. M.; Lawson, T.; Miller, M.; Thakkar, D. D.; Lord, J. W.; Yau, W. M.; Toth, B. 8-Arylguanine adducts from arenediazonium ions and DNA. *Chem. Biol. Interact.* **1996,** *101* (2), 149-164.

8. Gannett, P. M.; Shi, X.; Lawson, T.; Kolar, C.; Toth, B. Aryl radical formation during the metabolism of arylhydrazines by microsomes. *Chem. Res. Toxicol.* **1997,** *10* (12), 1372-1377.

9. Cadet, J.; Douki, T.; Gasparutto, D.; Ravanat, J. L. Oxidative damage to DNA: formation, measurement and biochemical features. *Mutat. Res.* **2003,** *531* (1-2), 5-23.

10. Kohen, R.; Nyska, A. Oxidation of biological systems: oxidative stress phenomena, antioxidants, redox reactions, and methods for their quantification. *Toxicol. Pathol.* **2002,** *30* (6), 620-650.

11. Martinez, G. R.; Loureiro, A. P.; Marques, S. A.; Miyamoto, S.; Yamaguchi, L. F.; Onuki, J.; Almeida, E. A.; Garcia, C. C.; Barbosa, L. F.; Medeiros, M. H.; Di, M. P. Oxidative and alkylating damage in DNA. *Mutat. Res.* **2003,** *544* (2-3), 115-127.

12. Cheng, K. C.; Cahill, D. S.; Kasai, H.; Nishimura, S.; Loeb, L. A. 8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G----T and A----C substitutions. *J. Biol. Chem.* **1992,** *267* (1), 166-172.

13. Runge-Morris, M.; Wu, N.; Novak, R. F. Hydrazine-mediated DNA damage: role of hemoprotein, electron transport, and organic free radicals. *Toxicol. Appl. Pharmacol.* **1994,** *125* (1), 123-132.

14. Gannett, P. M.; Powell, J. H.; Rao, R.; Shi, X.; Lawson, T.; Kolar, C.; Toth, B. C8-Arylguanine and C8-aryladenine formation in calf thymus DNA from arenediazonium ions. *Chem. Res. Toxicol.* **1999,** *12* (3), 297-304.

15. Kohda, K.; Tsunomoto, H.; Kasamatsu, T.; Sawamura, F.; Terashima, I.; Shibutani, S. Synthesis and miscoding specificity of oligodeoxynucleotide containing 8-phenyl-2'-deoxyguanosine. *Chem. Res. Toxicol.* **1997,** *10* (12), 1351-1358.

16. Hiramoto, K.; Kaku, M.; Kato, T.; Kikugawa, K. DNA strand breaking by the carbon-centered radical generated from 4-(hydroxymethyl) benzenediazonium salt, a carcinogen in mushroom Agaricus bisporus. *Chem. Biol. Interact.* **1995,** *94* (1), 21-36.

17. van Houte, L. P.; Westra, J. G.; Retel, J.; van, G. R. A spectroscopic study of the conformation of poly d(G-C).poly d(G-C) modified with the carcinogenic 2-aminofluorene. *Carcinogenesis* **1988,** *9* (6), 1017-1027.

18. Kimura, T.; Kawai, K.; Tojo, S.; Majima, T. One-electron attachment reaction of B- and Z-DNA modified by 8-bromo-2'-deoxyguanosine. *J. Org. Chem.* **2004,** *69* (4), 1169-1173.

19. Moller, A.; Nordheim, A.; Kozlowski, S. A.; Patel, D. J.; Rich, A. Bromination stabilizes poly(dG-dC) in the Z-DNA form under low-salt conditions. *Biochemistry* **1984,** *23* (1), 54-62.

20. Sugiyama, H.; Kawai, K.; Matsunaga, A.; Fujimoto, K.; Saito, I.; Robinson, H.; Wang, A. H. Synthesis, structure and thermodynamic properties of 8-methylguanine-containing oligonucleotides: Z-DNA under physiological salt conditions. *Nucleic Acids Res.* **1996,** *24* (7), 1272-1278.

21. Xu, Y.; Ikeda, R.; Sugiyama, H. 8-Methylguanosine: a powerful Z-DNA stabilizer. *J. Am. Chem. Soc.* **2003,** *125* (44), 13519-13524.

22. Gannett, P. M.; Heavner, S.; Daft, J. R.; Shaughnessy, K. H.; Epperson, J. D.; Greenbaum, N. L. Synthesis, properties, and NMR studies of a C8-phenylguanine modified oligonucleotide that preferentially adopts the Z DNA conformation. *Chem. Res. Toxicol.* **2003,** *16* (10), 1385-1394.

23. Wang, A. H.; Quigley, G. J.; Kolpak, F. J.; Crawford, J. L.; van Boom, J. H.; van der, M. G.; Rich, A. Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature* **1979,** *282* (5740), 680-686.

24. Rich, A.; Nordheim, A.; Wang, A. H. The chemistry and biology of left-handed Z-DNA. *Annu. Rev. Biochem.* **1984,** *53*, 791-846.

25. Sinden R.R. *DNA Structure and Function;* Academic Press, Inc.: San Diego, California, 1994.

26. Moller, A.; Nordheim, A.; Nichols, S. R.; Rich, A. 7-Methylguanine in poly(dG-dC).poly(dG-dC) facilitates z-DNA formation. *Proc. Natl. Acad. Sci. U. S. A* **1981,** *78* (8), 4777-4781.

27. Fujii, S.; Wang, A. H.; van der, M. G.; van Boom, J. H.; Rich, A. Molecular structure of (m5 dC-dG)3: the role of the methyl group on 5-methyl cytosine in stabilizing Z-DNA. *Nucleic Acids Res.* **1982,** *10* (23), 7879-7892.

28. Behe, M.; Felsenfeld, G. Effects of methylation on a synthetic polynucleotide: the B--Z transition in poly(dG-m5dC).poly(dG-m5dC). *Proc. Natl. Acad. Sci. U. S. A* **1981,** *78* (3), 1619-1623.

29. Peck, L. J.; Nordheim, A.; Rich, A.; Wang, J. C. Flipping of cloned d(pCpG)n.d(pCpG)n DNA sequences from right- to left-handed helical structure by salt, Co(III), or negative supercoiling. *Proc. Natl. Acad. Sci. U. S. A* **1982,** *79* (15), 4560-4564.

30. Nordheim, A.; Lafer, E. M.; Peck, L. J.; Wang, J. C.; Stollar, B. D.; Rich, A. Negatively supercoiled plasmids contain left-handed Z-DNA segments as detected by specific antibody binding. *Cell* **1982,** *31* (2 Pt 1), 309-318.

31. Rich, A.; Zhang, S. Timeline: Z-DNA: the long road to biological function. *Nat. Rev. Genet.* **2003,** *4* (7), 566-572.

32. Liu, L. F.; Wang, J. C. Supercoiling of the DNA template during transcription. *Proc. Natl. Acad. Sci. U. S. A* **1987,** *84* (20), 7024-7027.

33. Schroth, G. P.; Chou, P. J.; Ho, P. S. Mapping Z-DNA in the human genome. Computer-aided mapping reveals a nonrandom distribution of potential Z-DNA-forming sequences in human genes. *J. Biol. Chem.* **1992,** *267* (17), 11846-11855.

34. Champ, P. C.; Maurice, S.; Vargason, J. M.; Camp, T.; Ho, P. S. Distributions of Z-DNA and nuclear factor I in human chromosome 22: a model for coupled transcriptional regulation. *Nucleic Acids Res.* **2004,** *32* (22), 6501-6510.

35. Wittig, B.; Dorbic, T.; Rich, A. Transcription is associated with Z-DNA formation in metabolically active permeabilized mammalian cell nuclei. *Proc. Natl. Acad. Sci. U. S. A* **1991,** *88* (6), 2259-2263.

36. Wittig, B.; Wolfl, S.; Dorbic, T.; Vahrson, W.; Rich, A. Transcription of human c-myc in permeabilized nuclei is associated with formation of Z-DNA in three discrete regions of the gene. *EMBO J.* **1992,** *11* (12), 4653-4663.

37. Liu, R.; Liu, H.; Chen, X.; Kirby, M.; Brown, P. O.; Zhao, K. Regulation of CSF1 promoter by the SWI/SNF-like BAF complex. *Cell* **2001,** *106* (3), 309-318.

38. Garner, M. M.; Felsenfeld, G. Effect of Z-DNA on nucleosome placement. *J. Mol. Biol.* **1987,** *196* (3), 581-590.

39. Herbert, A. G.; Rich, A. A method to identify and characterize Z-DNA binding proteins using a linear oligodeoxynucleotide. *Nucleic Acids Res.* **1993,** *21* (11), 2669-2672.

40. Herbert, A. G.; Spitzner, J. R.; Lowenhaupt, K.; Rich, A. Z-DNA binding protein from chicken blood nuclei. *Proc. Natl. Acad. Sci. U. S. A* **1993,** *90* (8), 3339-3342.

41. Oh, D. B.; Kim, Y. G.; Rich, A. Z-DNA-binding proteins can act as potent effectors of gene expression in vivo. *Proc. Natl. Acad. Sci. U. S. A* **2002,** *99* (26), 16666-16671.

42. Schwartz, T.; Rould, M. A.; Lowenhaupt, K.; Herbert, A.; Rich, A. Crystal structure of the Zalpha domain of the human editing enzyme ADAR1 bound to left-handed Z-DNA. *Science* **1999,** *284* (5421), 1841-1845.

43. Freund, A. M.; Bichara, M.; Fuchs, R. P. Z-DNA-forming sequences are spontaneous deletion hot spots. *Proc. Natl. Acad. Sci. U. S. A* **1989,** *86* (19), 7465-7469.

44. Spitzner, J. R.; Chung, I. K.; Muller, M. T. Eukaryotic topoisomerase II preferentially cleaves alternating purine-pyrimidine repeats. *Nucleic Acids Res.* **1990,** *18* (1), 1-11.

45. Wang, G.; Christensen, L. A.; Vasquez, K. M. Z-DNA-forming sequences generate large-scale deletions in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A* **2006,** *103* (8), 2677-2682.

46. Lawson, T.; Gannett, P.; Yau, W.; Dalal N.S.; Toth, B. Different Patterns of Mutagenicity of Arenediazonium Ion in V79 Cells and Samonella typhimurium TA102: Evidence for Different Mechanisms of Action. *J. Agric. Food Chem.* **1995,** *43*, 2627-2635.

47. Powell, J. H.; Gannett, P. M. Mechanisms of carcinogenicity of aryl hydrazines, aryl hydrazides, and arenediazonium ions. *J. Environ. Pathol. Toxicol. Oncol.* **2002,** *21* (1), 1-31.

48. RamaKrishna, N. V.; Gao, F.; Padmavathi, N. S.; Cavalieri, E. L.; Rogan, E. G.; Cerny, R. L.; Gross, M. L. Model adducts of benzo[a]pyrene and nucleosides formed from its radical cation and diol epoxide. *Chem. Res. Toxicol.* **1992,** *5* (2), 293-302.

49. Chen, L.; Devanesan, P. D.; Higginbotham, S.; Ariese, F.; Jankowiak, R.; Small, G. J.; Rogan, E. G.; Cavalieri, E. L. Expanded analysis of benzo[a]pyrene-DNA adducts formed in vitro and in mouse skin: their significance in tumor initiation. *Chem. Res. Toxicol.* **1996,** *9* (5), 897-903.

50. Rogan, E. G.; Devanesan, P. D.; RamaKrishna, N. V.; Higginbotham, S.; Padmavathi, N. S.; Chapman, K.; Cavalieri, E. L.; Jeong, H.; Jankowiak, R.; Small, G. J. Identification and quantitation of benzo[a]pyrene-DNA adducts formed in mouse skin. *Chem. Res. Toxicol.* **1993,** *6* (3), 356-363.

51. Loeb, L. A.; Preston, B. D. Mutagenesis by apurinic/apyrimidinic sites. *Annu. Rev. Genet.* **1986,** *20*, 201-230.

52. Mol, C. D.; Hosfield, D. J.; Tainer, J. A. Abasic site recognition by two apurinic/apyrimidinic endonuclease families in DNA base excision repair: the 3' ends justify the means. *Mutat. Res.* **2000,** *460* (3-4), 211-229.

53. Pogozelski, W. K.; Tullius, T. D. Oxidative Strand Scission of Nucleic Acids: Routes Initiated by Hydrogen Abstraction from the Sugar Moiety. *Chem. Rev.* **1998,** *98*, 1089-1107.

54. Boiteux, S.; Laval, J. Coding properties of poly(deoxycytidylic acid) templates containing uracil or apyrimidinic sites: in vitro modulation of mutagenesis by deoxyribonucleic acid repair enzymes. *Biochemistry* **1982,** *21* (26), 6746-6751.

55. Dutta, S.; Chowdhury, G.; Gates, K. S. Interstrand cross-links generated by abasic sites in duplex DNA. *J. Am. Chem. Soc.* **2007,** *129* (7), 1852-1853.

56. Applied Biosystems Models 392 and 394 DNA/RNA Synthesizers User's Manual. 1994.

57. Adams, S. P.; Kavka, K. S.; Wykes, E. J.; Holder, S. B.; Gulluppi, G. R. Hindered Dialkylamino Nucleoside Phosphite Reagents in the Synthesis of Two DNA 51-Mers. *J. Am. Chem. Soc.* **1983,** *105*, 661-663.

58. Agrofoglio, L. A.; Gillaizeau, I.; Saito, Y. Palladium-Assisted Routes to Nucleosides. *Chem. Rev.* **2003,** *103*, 1875-1916.

59. Hocek, M. Syntheses of Purines Bearing Carbon Substituents in Positions 2, 6 or 8 by Metal- or Organometal-Mediated C-C Bond-Forming Reactions. *Eur. J. Org. Chem.* **2003,** 245-254.

60. Amann, N.; Wagenknecht, H. Preperation of Pyrenyl-Modified Nucleosides via Suzuki-Miyaura Cross-Coupling Reactions. *Synlett* **2002,** (5), 687-691.

61. Western, E. C.; Daft, J. R.; Johnson, E. M. I.; Gannett, P. M.; Shaughnessy, K. H. Efficient One-Step Suzuki Arylation of Unprotected Halonucleosides, Using Water-Soluble Palladium Catalysts. *J. Org. Chem.* **2003,** *68*, 6767-6774.

62. Cepanec, I. *Synthesis of Biaryls;* Elsevier Ltd.: Oxford, UK, 2004.

63. Miyaura, N.; Yanagi, T.; Suzuki, A. The Palladium-Catalyzed Cross-Coupling Reaction of Phenylboronic Acid with Haloarenes in the Presence of Bases. *Syn. Commun.* **1981,** *11* (7), 513-519.

64. Miyaura, N.; Yamada, K.; Suginome, H.; Suzuki, A. Novel and Convenient Method for the Stereo- and Regiospecific Synthesis of Conjugated Alkadienes and Alkenynes via the Palladium-Catalyzed Cross-Coupling Reaction of 1-Alkenylboranes with Bromoalkenes and Bromoalkynes. *J. Am. Chem. Soc.* **1985,** *107*, 972-980.

65. Suzuki, A. Organoboron compounds in new synthetic reactions. *Pure & Appl. Chem.* **1995,** *57* (12), 1749-1758.

66. Miyaura, N.; Suzuki, A. Palladium-Catalyzed Cross-Coupling Reactions of Organoboron Compounds. *Chem. Rev.* **1995,** *95*, 2457-2483.

67. Shaughnessy, K. H.; Booth, R. S. Strically Demanding, Water-Soluble Alkylphosphines as Ligands for High Activity Suzuki Coupling of Aryl Bromides in Aqueous Solvents. *Org. Lett.* **2001,** *3* (17), 2757-2759.

68. Sonveaux, E. Protecting Groups in Oligonucleotide Synthesis. In *Protocols for Oligonucleotide Conjugates*, Agrawal, S., Ed.; Humana Press: Totowa, New Jersey, 1994.

69. Hawthrone, M. F.; Lewis, E. S. Amine boranes III: Hydrolysis of pyridine diphenylborane and the mechanism of hydride transfer reactions. *J. Am. Chem. Soc.* **1958,** *80*, 4296-4299.

70. Zheng, N.; Armstrong, J. D.; Eng, K. K.; et al. A convergent asymmetric synthesis of a growth hormone secretagogue. *Tetrahedron: Asymmetry* **2003,** *14*, 3435-3446.

71. Koisume, S.; Kamiya, H.; Inoue, H.; Ohtsuka, E. Synthesis and thermodynamic stabilities of damaged DNA involving 8-hydroxyguanine (7,8-dihydro-8-oxoguanine) in a ras-gene fragment. *Nucleosides & Nucleotides* **1994,** *13* (6&7), 1517-1534.

72. Fabrega, C.; Macias, M. J.; Eritja, R. Synthesis and properties of oligonucleotides containing 8-bromo-2'-deoxyguanosine. *Nucleosides, Nucleotides & Nucleic acids* **2001,** *20* (3), 251-260.

73. Daft, J. R. Arylhydrazine Carcinogenesis and the Synthesis of C8-Arylpurine Oligonucleotides: A Study of DNA Adduct Effects on DNA Conformation and Stability. Department of Basic Pharmaceutical Sciences, West Virginia University, 2005.

74. Mclaughlin, L. W. Mixed-mode chromatography of nucleic acids. *Chem. Rev.* **1989,** *89*, 309-319.

75. Smith, R. M. *Understanding Mass Spectra: A Basic Approach;* 2nd ed.; John Wiley & Sons, Inc.: Hoboken, New Jersey, 2004.

76. Beck, J. L.; Colgrave, M. L.; Ralph, S. F.; Sheil, M. M. Electrospray ionization mass spectrometry of oligonucleotide complexes with drugs, metals, and proteins. *Mass Spectrometry Reviews* **2001,** *20*, 61-87.

77. Banoub, J. H.; Newton, R. P.; Esmans, E.; Ewing, D. F.; Mackenzie, G. Recent developments in mass spectrometry for the characterization of nucleosides, nucleotides, oligonucleotides, and nucleic Acids. *Chem. Rev.* **2005,** *105*, 1869-1915.

78. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **1986,** *246*, 64-71.

79. Emyth, E. T. *Finnigan LCQ Deca: Getting Start;* Revision A ed.; Technical Publications, ThermoQuest Corporation: San Jose, California, 1999.

80. Gaskell, S. J. Electrospray: Principles and practice. *J. Mass Spectrometry* **1997,** *32*, 677-688.

81. Johnson, W. C. CD of Nucleic Acids. In *Circular Dichroism: Principles and Applications*, 2nd ed.; Berova, N., Nakanishi, K., Woody R.W., Eds.; John Wiley & Sons, Inc.: New York, NY, 2000; pp 703-718.

82. Pohl, F. M.; Jovin, T. M. Salt-induced co-operative conformational change of a synthetic DNA: equilibrium and kinetic studies with poly (dG-dC). *J. Mol. Biol.* **1972,** *67* (3), 375-396.

83. Albergo, D. D.; Turner, D. H. Solvent effects on thermodynamics of double-helix formation in (dG-dC)3. *Biochemistry* **1981,** *20* (6), 1413-1418.

84. Zimmer, C.; Tymen, S.; Marck, C.; Guschlbauer, W. Conformational transitions of poly(dA-dC).poly(dG-dT) induced by high salt or in ethanolic solution. *Nucleic Acids Res.* **1982,** *10* (3), 1081-1091.

85. Xodo, L. E.; Manzini, G.; Quadrifoglio, F.; van der Marel, G. A.; van Boom, J. H. The B-Z conformational transition in folded oligodeoxynucleotides: loop size and stability of Z-hairpins. *Biochemistry* **1988,** *27* (17), 6327-6331.

86. Silverstein, R. M.; Webster, F. X.; Kiemle, D. J. *Spectrometric Identification of Organic Compounds;* 7th ed.; John Wiley & Sons, Inc.: Hoboken, New Jersey, 2005.

87. Nelson, J. H. *Nuclear Magnetic Resonance Spectroscopy;* Pearson Education, Inc.: Upper Saddle River, New Jersey, 2003.

88. Crews, P.; Rodriguez, J.; Jarpars, M. *Organic Structure Analysis;* Oxford University Press, Inc.: New York, New York, 1998.

89. Bachers, G. E.; Schaefer, T. Applications of the intramolecular nuclear Overhauser effect in structural organic chemistry. *Chem. Rev.* **1971,** *71* (6), 617-626.

90. Solomon, I. Relaxation processes in a system of two spins. *Phys. Rev.* **1955,** *99* (2), 559-566.

91. Patel, D. J.; Canuel, L. L.; Pohl, F. M. "Alternating B-DNA" conformation for the oligo(dG-dC) duplex in high-salt solution. *Proc. Natl. Acad. Sci. U. S. A* **1979,** *76* (6), 2508-2511.

92. Scheek, R. M.; Boelens, R.; Russo, N.; van Boom, J. H.; Kaptein, R. Sequential resonance assignments in 1H NMR spectra of oligonucleotides by two-dimensional NMR spectroscopy. *Biochemistry* **1984,** *23* (7), 1371-1376.

93. Mitra, C. K.; Sarma, M. H.; Sarma, R. H. Left-handed deoxyribonucleic acid double helix in solution. *Biochemistry* **1981,** *20* (7), 2036-2041.

94. Feigon, J.; Wang, A. H.; van der Marel, G. A.; van Boom, J. H.; Rich, A. A one- and two-dimensional NMR study of the B to Z transition of (m5dC-dG)3 in methanolic solution. *Nucleic Acids Res.* **1984,** *12* (2), 1243-1263.

95. Orbons, L. P.; van der Marel, G. A.; van Boom, J. H.; Altona, C. The B and Z forms of the d(m5C-G)3 and d(br5C-G)3 hexamers in solution. A 300-MHz and 500-MHz two-dimensional NMR study. *Eur. J. Biochem.* **1986,** *160* (1), 131-139.

96. Ikuta, S.; Wang, Y. S. Conformation and dynamics of Z-DNA oligomer duplex of d[(CG)3TATA(CG)3] in solution. *Nucleic Acids Res.* **1989,** *17* (11), 4131-4144.

97. Virgilio, A.; Esposito, V.; Randazzo, A.; Mayol, L.; Galeone, A. 8-methyl-2'-deoxyguanosine incorporation into parallel DNA quadruplex structures. *Nucleic Acids Res.* **2005,** *33* (19), 6188-6195.

98. Herbert, A.; Rich, A. Left-handed Z-DNA: structure and function. *Genetica* **1999,** *106* (1-2), 37-47.

99. Heavner, S.; Gannett, P. M. Molecular dynamics and free energy calculations of the B and Z forms of C8-arylguanine modified oligonucleotides. *J. Biomol. Struct. Dyn.* **2005,** *23* (2), 203-220.

100. Lavery, R.; Zakrzewska, K. Base and Base Pair Morphologies, Helical Parameters, and Definitions. In *Oxford Handbook of Nucleic Acid Structure*, Neidle, S., Ed.; Oxford University Press: New York, New York, 1999; pp 39-76.

101. Heavner, S. Molecular Modeling and Experimental Determination of the Structure of C8-Arylguanine Modified Oligonucleotides that Preferentially Adopt the Z-DNA Conformation. Department of Basic Pharmaceutical Sciences, West Virginia University, 2004.

102. Case, D. A.; Darden, T. A.; Cheatham, I. T. E.; et al *AMBER 8;* University of California, San Francisco: San Francisco, California, 2004.

103. Lavery, R.; Sklenar, H. The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.* **1988,** *6* (1), 63-91.

104. *CURVES,* version 5.2; Laboratoire de Biochimie Théorique URA 77 CNRS, Institut de Biologie Physico-Chimique: Paris, France, 1997

105. *3DNA,* version 1.5; Department of Chemistry, Rutgers University: Piscataway, New Jersey, 2002

106. *MOIL-View,* version 10.2; Department of Pharmaceutical Chemistry, University of California, San Francisco: San Francisco, California, 1999

107. Naim, M.; Bhat, S.; Rankin, K. N.; Dennis, S.; Chowdhury, S. F.; Siddiqi, I.; Drabik, P.; Sulea, T.; Bayly, C. I.; Jakalian, A.; Purisima, E. O. Solvated interaction energy (SIE) for scoring protein-ligand binding affinities. 1. Exploring the parameter space. *J. Chem. Inf. Model.* **2007,** *47* (1), 122-133.

108. Cui, Q.; Sulea, T.; Schrag, J. D.; Munger, C.; Hung, M. N.; Naim, M.; Cygler, M.; Purisima, E. O. Molecular dynamics-solvated interaction energy studies of protein-protein interactions: the MP1-p14 scaffolding complex. *J. Mol. Biol.* **2008,** *379* (4), 787-802.

109. *GaussView,* version 2; Gaussian, Inc.: Pittsburgh, Pennsylvania, 1998

110. Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; et al. *AMBER 7;* University of California, San Francisco: San Francisco, California, 2002.

111. Purisima, E. O.; Nilar, S. H. A simple yet accurate boundary element method for continuum dielectric calculations. *J. Comput. Chem.* **1995,** *16*, 681-689.

112. Purisima, E. O. Fast summation boundary element method for calculating solvation free energies of macromolecules. *J. Comput. Chem.* **1998,** *19*, 1494-1504.

113. Levin, D. E.; Hollstein, M.; Christman, M. F.; Schwiers, E. A.; Ames, B. N. A new Salmonella tester strain (TA102) with A X T base pairs at the site of

mutation detects oxidative mutagens. *Proc. Natl. Acad. Sci. U. S. A* **1982,** *79* (23), 7445-7449.

114. Van Zeeland, A. A.; Simons, J. W. Linear dose--response relationships after prolonged expression times in V-79 Chinese hamster cells. *Mutat. Res.* **1976,** *35* (1), 129-137.

115. Bradley, M. O.; Bhuyan, B.; Francis, M. C.; Langenbach, R.; Peterson, A.; Huberman, E. Mutagenesis by chemical agents in V79 chinese hamster cells: a review and analysis of the literature. A report of the Gene-Tox Program. *Mutat. Res.* **1981,** *87* (2), 81-142.

116. Pegg, A. E.; Hui, G. Formation and subsequent removal of O6-methylguanine from deoxyribonucleic acid in rat liver and kidney after small doses of dimethylnitrosamine. *Biochem. J.* **1978,** *173* (3), 739-748.

117. Herron, D. C.; Shank, R. C. In vivo kinetics of O6-methylguanine and 7-methylguanine formation and persistence in DNA of rats treated with symmetrical dimethylhydrazine. *Cancer Res.* **1981,** *41* (10), 3967-3972.

118. Lafer, E. M.; Sousa, R.; Ali, R.; Rich, A.; Stollar, B. D. The effect of anti-Z-DNA antibodies on the B-DNA-Z-DNA equilibrium. *J. Biol. Chem.* **1986,** *261* (14), 6438-6443.

119. Thomas, T. J.; Seibold, J. R.; Adams, L. E.; Hess, E. V. Hydralazine induces Z-DNA conformation in a polynucleotide and elicits anti(Z-DNA) antibodies in treated patients. *Biochem. J.* **1993,** *294 (Pt 2)*, 419-425.

120. Wolfl, S.; Martinez, C.; Rich, A.; Majzoub, J. A. Transcription of the human corticotropin-releasing hormone gene in NPLC cells is correlated with Z-DNA formation. *Proc. Natl. Acad. Sci. U. S. A* **1996,** *93* (8), 3664-3668.

121. Runkel, L.; Nordheim, A. Chemical footprinting of the interaction between left-handed Z-DNA and anti-Z-DNA antibodies by diethylpyrocarbonate carbethoxylation. *J. Mol. Biol.* **1986,** *189* (3), 487-501.

122. Oyoshi, T.; Kawai, K.; Sugiyama, H. Efficient C2'alpha-hydroxylation of deoxyribose in protein-induced Z-form DNA. *J. Am. Chem. Soc.* **2003,** *125* (6), 1526-1531.

123. Kim, Y. G.; Muralinath, M.; Brandt, T.; Pearcy, M.; Hauns, K.; Lowenhaupt, K.; Jacobs, B. L.; Rich, A. A role for Z-DNA binding in vaccinia virus pathogenesis. *Proc. Natl. Acad. Sci. U. S. A* **2003,** *100* (12), 6974-6979.

124. Kahmann, J. D.; Wecking, D. A.; Putter, V.; Lowenhaupt, K.; Kim, Y. G.; Schmieder, P.; Oschkinat, H.; Rich, A.; Schade, M. The solution structure of the N-terminal domain of E3L shows a tyrosine conformation that may explain its

reduced affinity to Z-DNA in vitro. *Proc. Natl. Acad. Sci. U. S. A* **2004,** *101* (9), 2712-2717.

125. Quyen, D. V.; Ha, S. C.; Lowenhaupt, K.; Rich, A.; Kim, K. K.; Kim, Y. G. Characterization of DNA-binding activity of Z alpha domains from poxviruses and the importance of the beta-wing regions in converting B-DNA to Z-DNA. *Nucleic Acids Res.* **2007,** *35* (22), 7714-7720.

APPENDIX A

ABBREVIATIONS

Chemicals

8-BrdG, 8-bromo-2′-deoxyguanosine
A, adenine
ACN, acetonitrile
BD, benzenediazonium ion
C, cytosine
CBD, *p*-carboxybenzenediazonium ion
CE-PA, cyanoethyl phosphoramidites
CG, $d(CGCGCGCGCG)_2$
CG^{8CPh}, $d(CGCGCG^{8CPh}CGCG)_2$
CG^{8HMPh}, $d(CGCGCG^{8HMPh}CGCG)_2$
CG^{8MMPh}, $d(CGCGCG^{8MMPh}CGCG)_2$
CG^{8Ph}, $d(CGCGCG^{8Ph}CGCG)_2$
CG^{8Tol}, $d(CGCGCG^{8Tol}CGCG)_2$
D2O, deuterium oxide
DEME, diethylaminoethyl
dG, 2′-deoxyguanosine
DME, dimethylether
DMF, N'N-dimethylformamide
DMTr, dimethoxytrityl
dsDNA, double strand DNA
G, guanine
HMBD, *p*-hydroxymethylbenzenediazonium ions
MBD, *p*-methylbenzenediazonium ions
MMBD, *p*-methoxymethylbenzenediazonium ions
NBS, N-bromosuccinimide
$Pd(OAc)_2$, palladium acetate
ssDNA, single strand DNA
T, thymine
TBS, t-butyldimethylsilyl
TCA, trichloroacetic acid

Experimentals

1D, one dimension
2D, two dimensions
COSY, correlation spectroscopy
C_{ss}, conc. of ssDNA (μM)
C_{ds}, conc. of dsDNA (μM)
C_B, conc. of B-DNA (μM)
C_Z, conc. of Z-DNA (μM)
CD, circular dichroism
CD^{295}, CD at 295 nm
CI, chemical ionization
CPG, controlled pore glass
CSF-1, colony stimulating factor-1
EI, electron ionization
ESI, electrospray ionization
FEP, free energy perturbation
FPLC, fast protein liquid chromatography
HETCOR, heteronuclear correlation spectroscopy
HPLC, high pressure liquid chromatography
M, molecular mass
MD, molecular dynamic
MM, molecular mechanic
MS, mass Spectrometry
MW, molecular weight
m/z, mass to charge ratio
NMR, nuclear magnetic resonance
nOe, nuclear Overhauser enhancement
NOESY, nuclear Overhauser effect correlation spectroscopy
N_α, lower energy state nuclei
N_β, higher energy state nuclei
pD, apparent pD
R, gas constant
RF, radio frequency
RMSD, root mean square derivation
RP-TLC, reverse phase thin layer chromatography
SPE, solid phase extraction
TLC, thin layer chromatography
Tm, melting temperature
UV, ultraviolet
ΔE, energy gap
ΔG, free energy
ΔH, enthalpy
ΔS, entropy
A, absorbance
B_0, magnetic field strength
b, path length
f_{ss}, molar fraction of ssDNA
f_{ds}, molar fraction of dsDNA
f_B, molar fraction of B-DNA
f_Z, molar fraction of Z-DNA
h, Plank's constant
I, spin number
J, coupling constant
n, charge state of the molecule (MS); number of adjacent protons with the same coupling constant (NMR)
ν, applied radio frequency
ν_{eff}, effective radio frequency
σ, shielding constant
δ, chemical shift
ε, molar absorptivity
ϵ, total ellipticity
Δϵ, molar ellipticity
γ, gyromagnetic ratio

APPENDIX B
CHEMICAL STRUCTURES

1) 8-Bromo-2′-deoxyguanosine

2) 4-(TBS-O-methyl)-phenyl bromide

3) 4-(TBS-O-methyl)-phenyl boronic acid

4) 8-(4-Carboxyphenyl)-2′-deoxyguanosine

5) 8-(4-Methoxymethylphenyl)-2′-deoxyguanosine

6) 8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine

7) N2-(N,N-Dimethylformamidine)-8-(4-carboxyphenyl)-2′-deoxyguanosine

8) N2-(N,N-Dimethylformamidine)-8-(4-methoxymethylphenyl)-2′-deoxyguanosine

9) N2-(N,N-Dimethylformamidine)-8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine

10) 5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-carboxyphenyl)-2′-deoxyguanosine

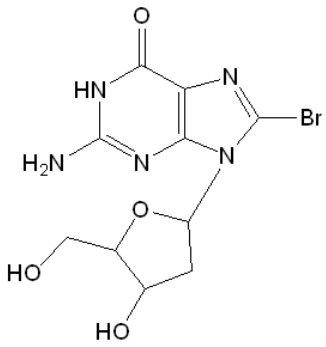
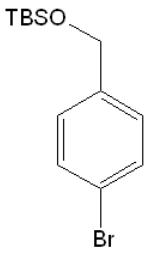
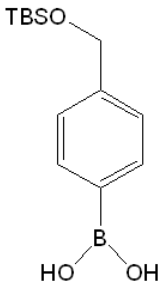
11) 5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-methoxymethylphenyl)-2′-deoxyguanosine

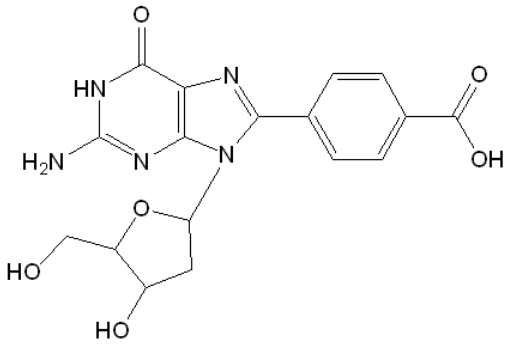
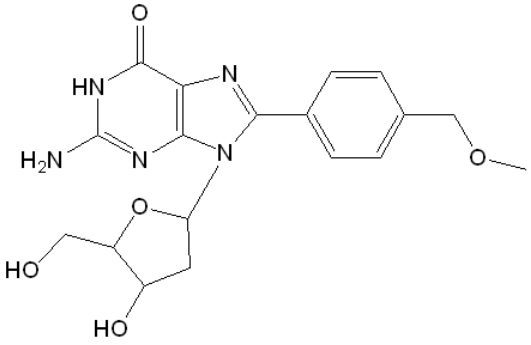
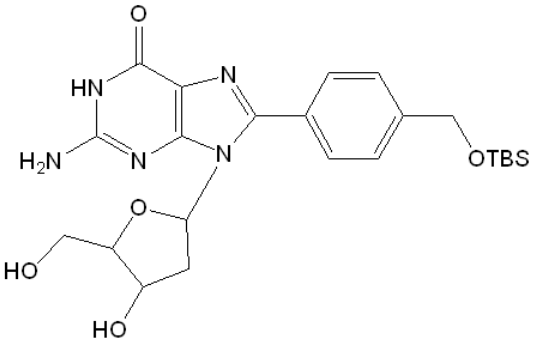
12) 5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine

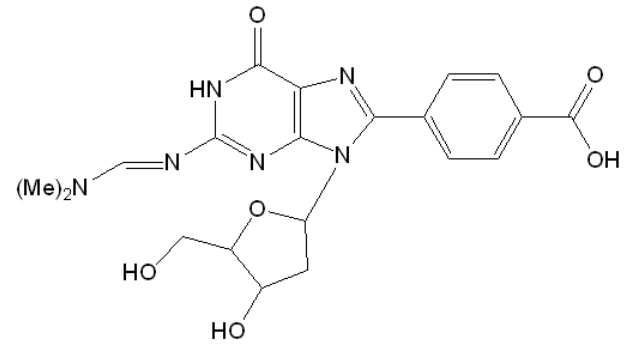
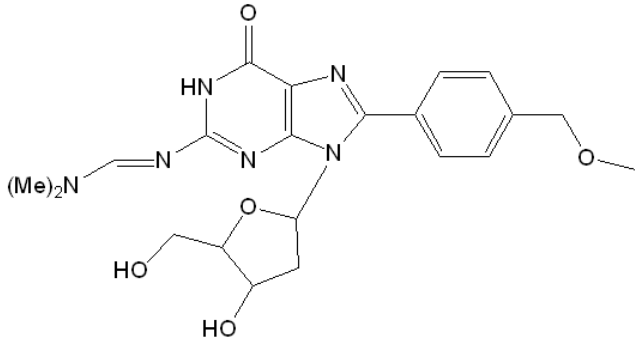
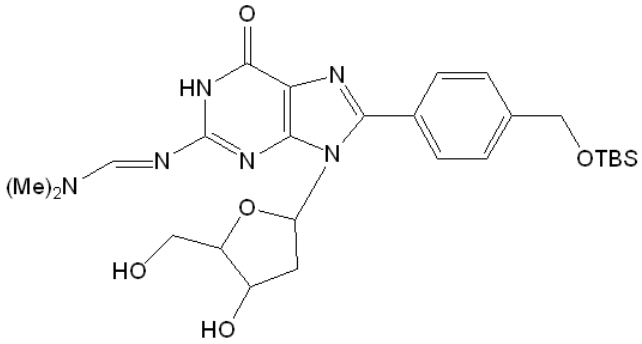
13) 3′-O-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-carboxyphenyl)-2′-deoxyguanosine

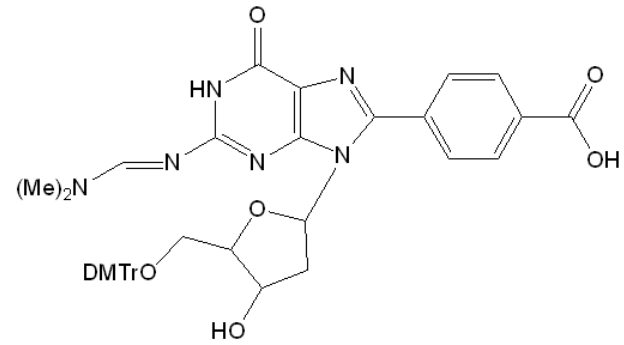
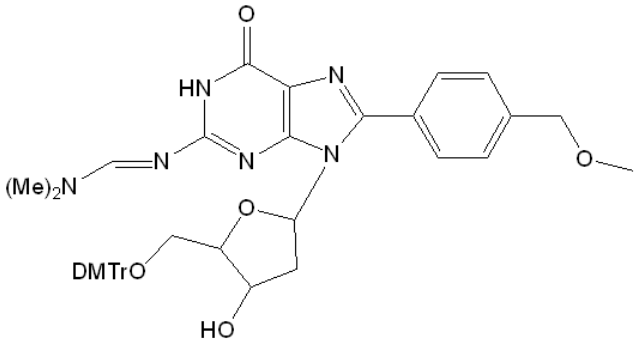
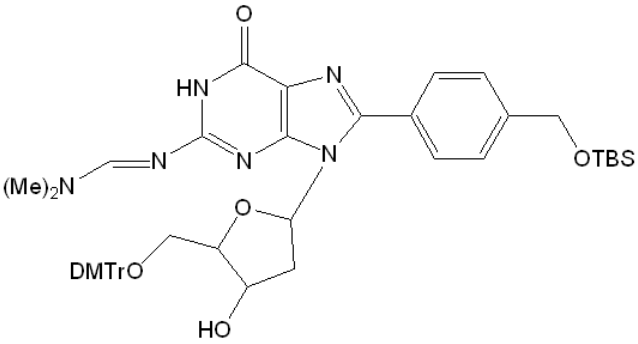
14) 3′-O-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-methoxymethylaryl)-2′-deoxyguanosine

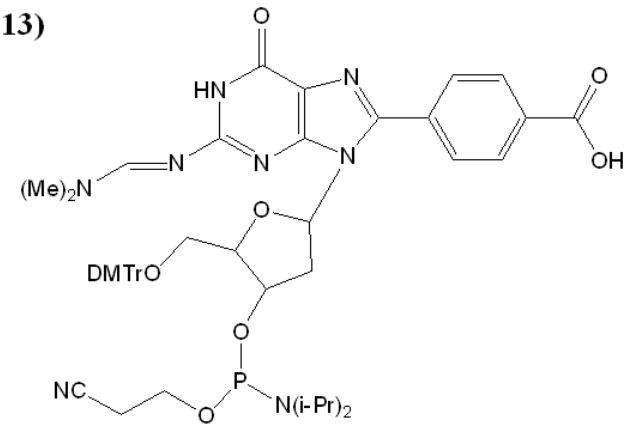
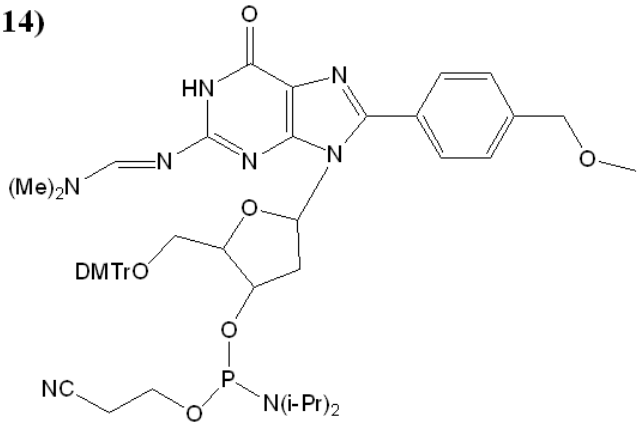
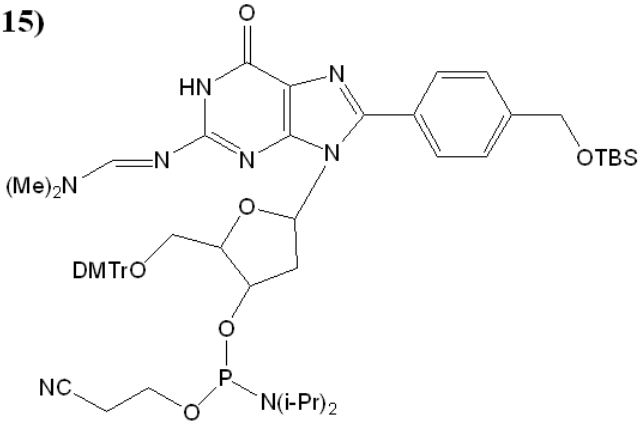
15) 3′-O-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-(TBS-O-methyl)aryl)-2′-deoxyguanosine

Chemicals	Formular	MW
1) 	$C_{10}H_{12}O_4N_5Br$	346.14
2) 	$C_{13}H_{21}OSiBr$	301.30
3) 	$C_{13}H_{23}O_3BSi$	266.22

Chemicals	Formular	MW
4) 	$C_{17}H_{17}O_6N_5$	387.35
5) 	$C_{18}H_{21}O_5N_5$	387.40
6) 	$C_{23}H_{33}O_5N_5Si$	487.63

Chemicals	Formular	MW
7) 	$C_{20}H_{22}O_6N_6$	442.43
8) 	$C_{21}H_{26}O_5N_6$	442.47
9) 	$C_{26}H_{38}O_5N_6Si$	542.71

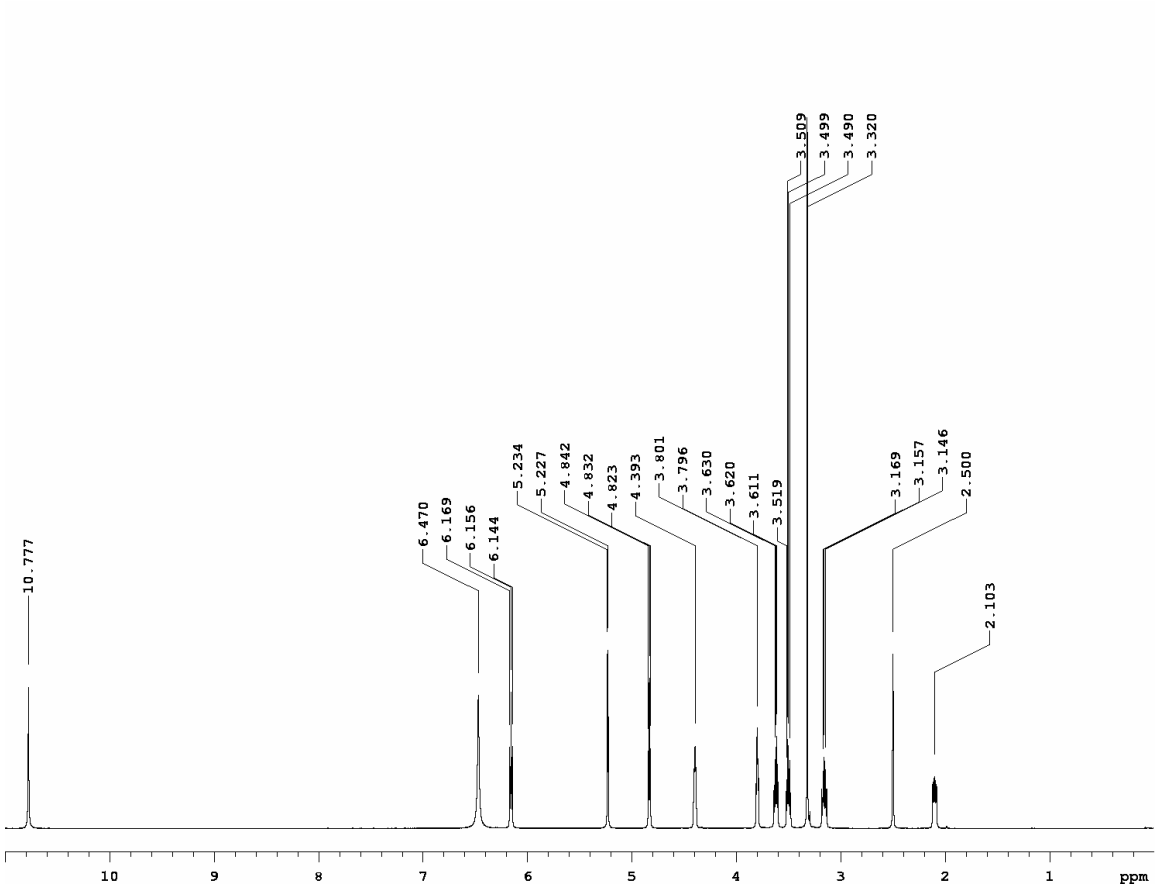
Chemicals	Formular	MW
10) 	$C_{41}H_{40}O_8N_6$	744.80
11) 	$C_{42}H_{44}O_7N_6$	744.85
12) 	$C_{47}H_{56}O_7N_6Si$	845.08

Chemicals	Formular	MW
13) 	$C_{50}H_{57}O_9N_8P$	945.02
14) 	$C_{51}H_{61}O_8N_8P$	945.07
15) 	$C_{56}H_{73}O_8N_8SiP$	1045.30

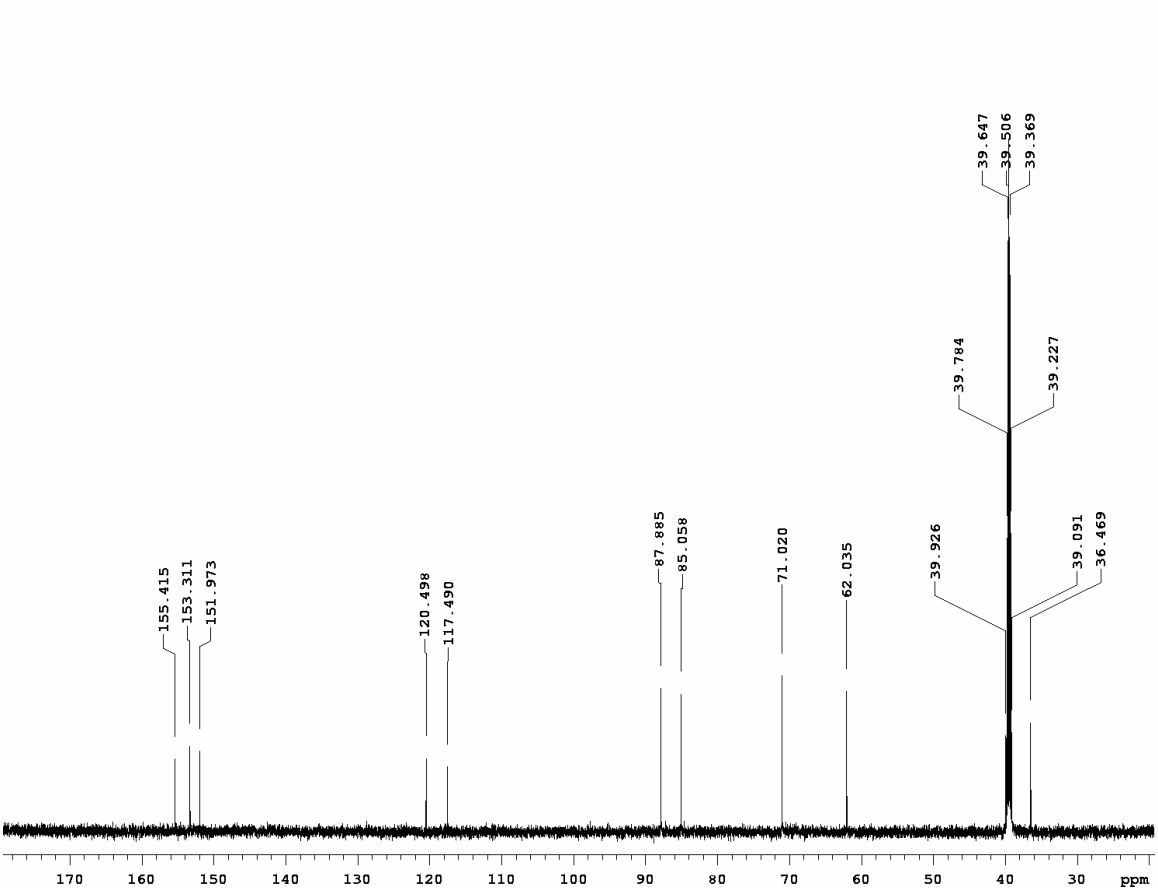
APPENDIX C

NMR, UV, and MS SPECTRA

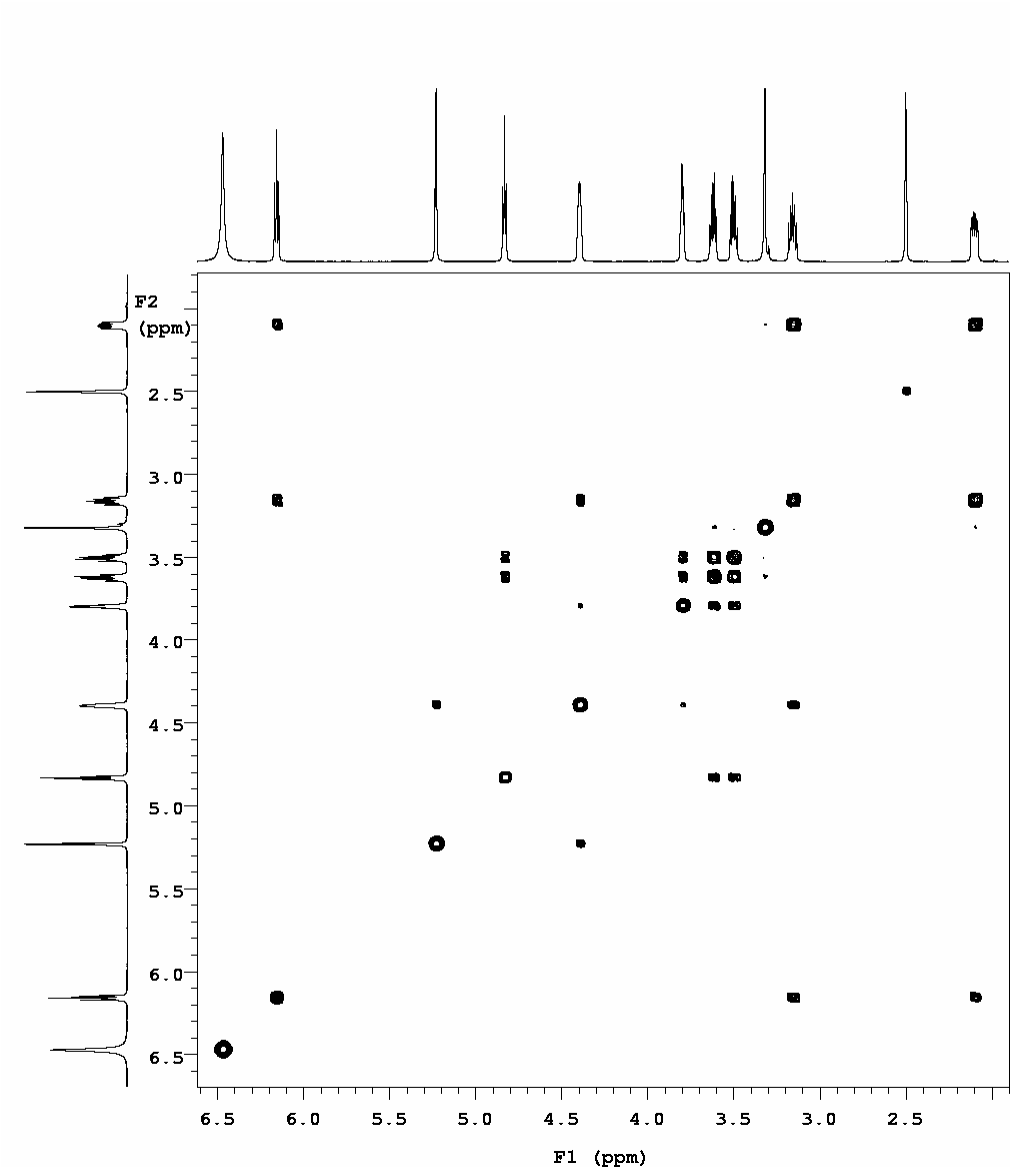
^{1}H NMR of 8-Bromo-2′-deoxyguanosine (1)



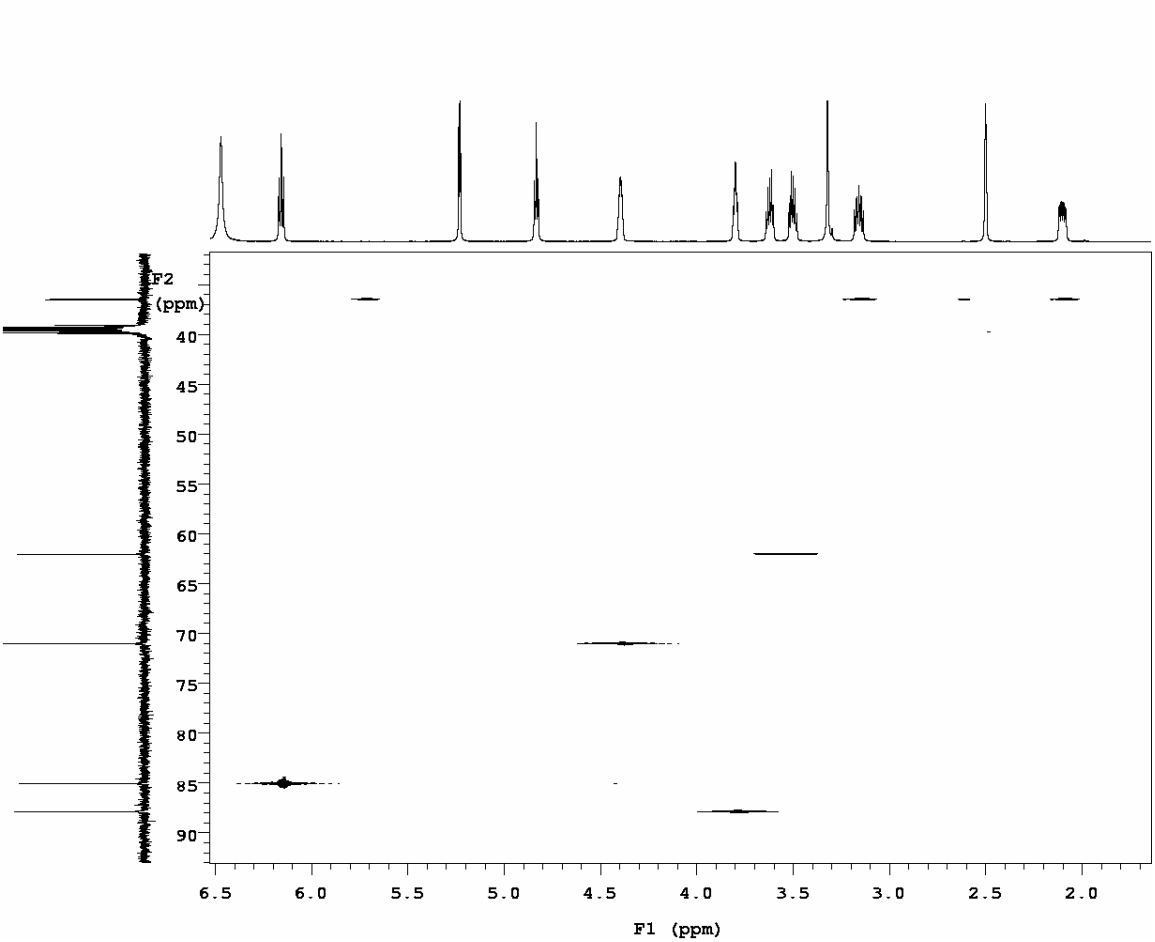
^{13}C NMR of 8-Bromo-2′-deoxyguanosine (1)



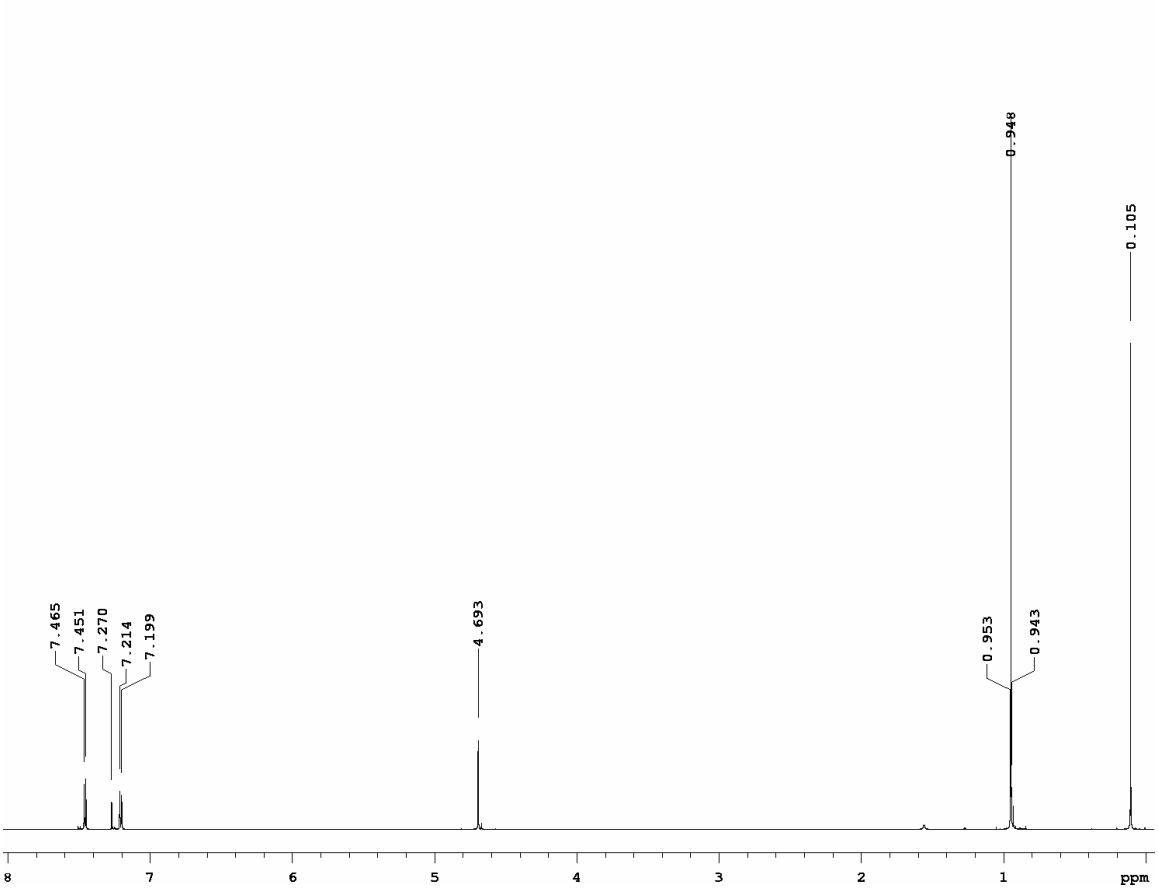
^{1}H-^{1}H COSY NMR of 8-Bromo-2′-deoxyguanosine (1)



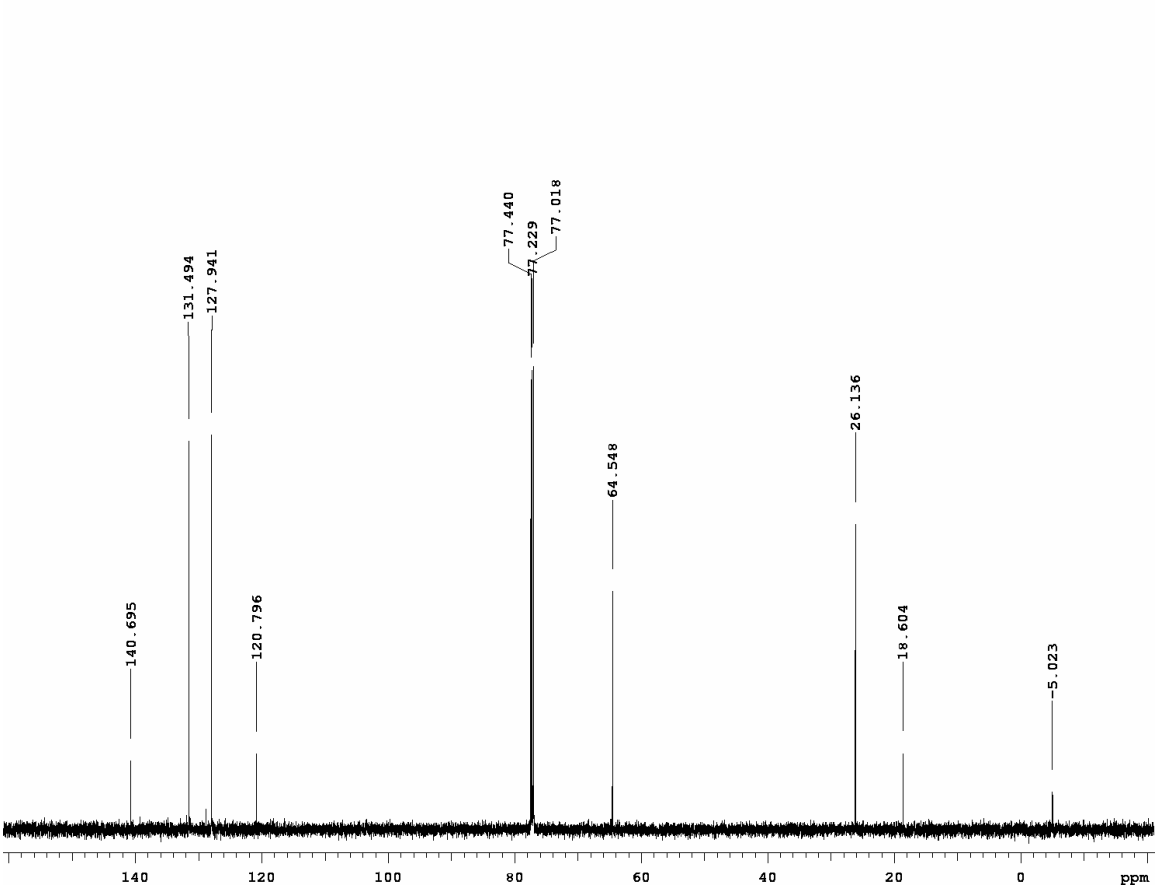
^{1}H-^{13}C HETCOR NMR of 8-Bromo-2′-deoxyguanosine (1)



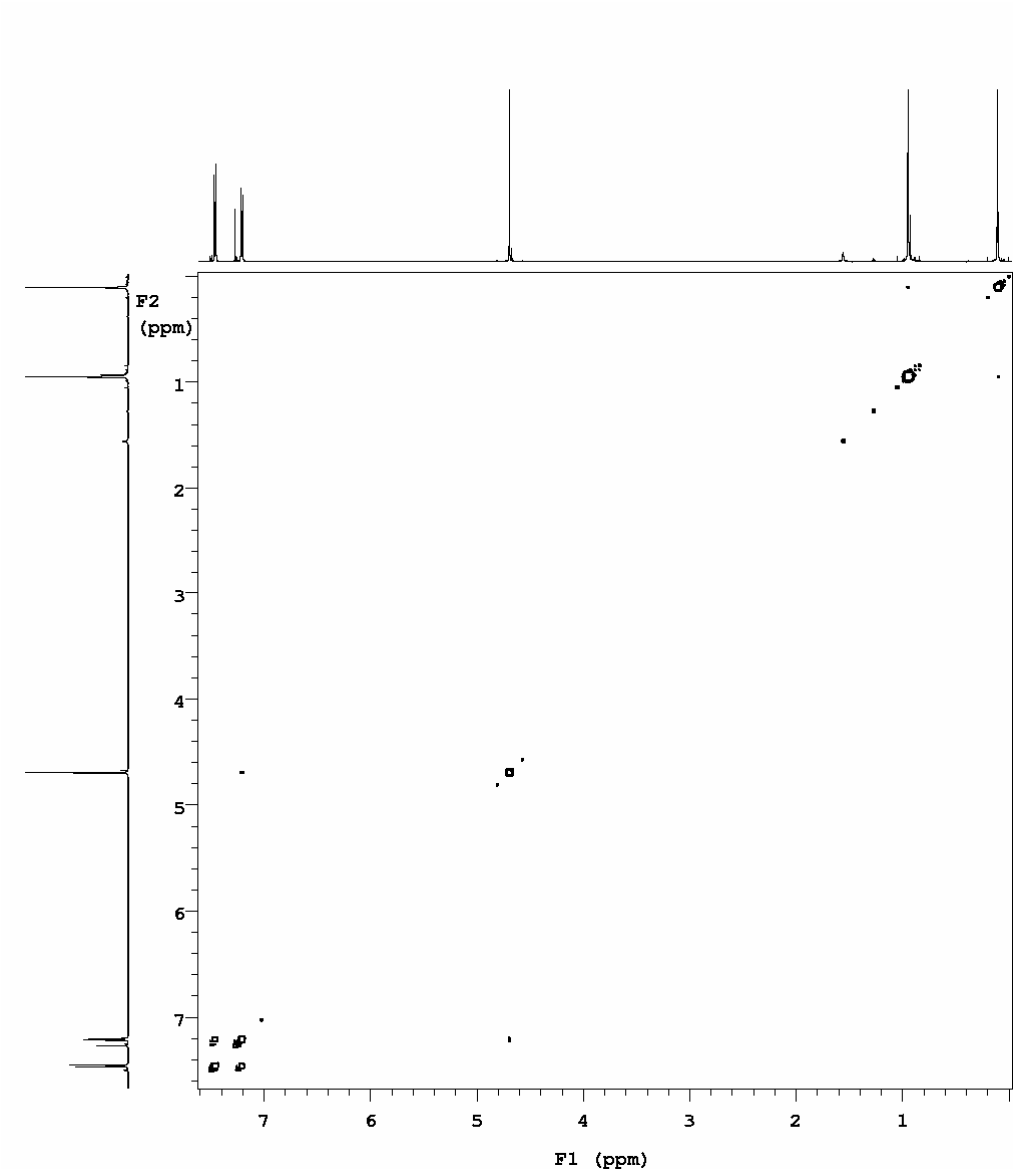
^{1}H NMR of 4-(TBS-O-methyl)-phenyl bromide (2)



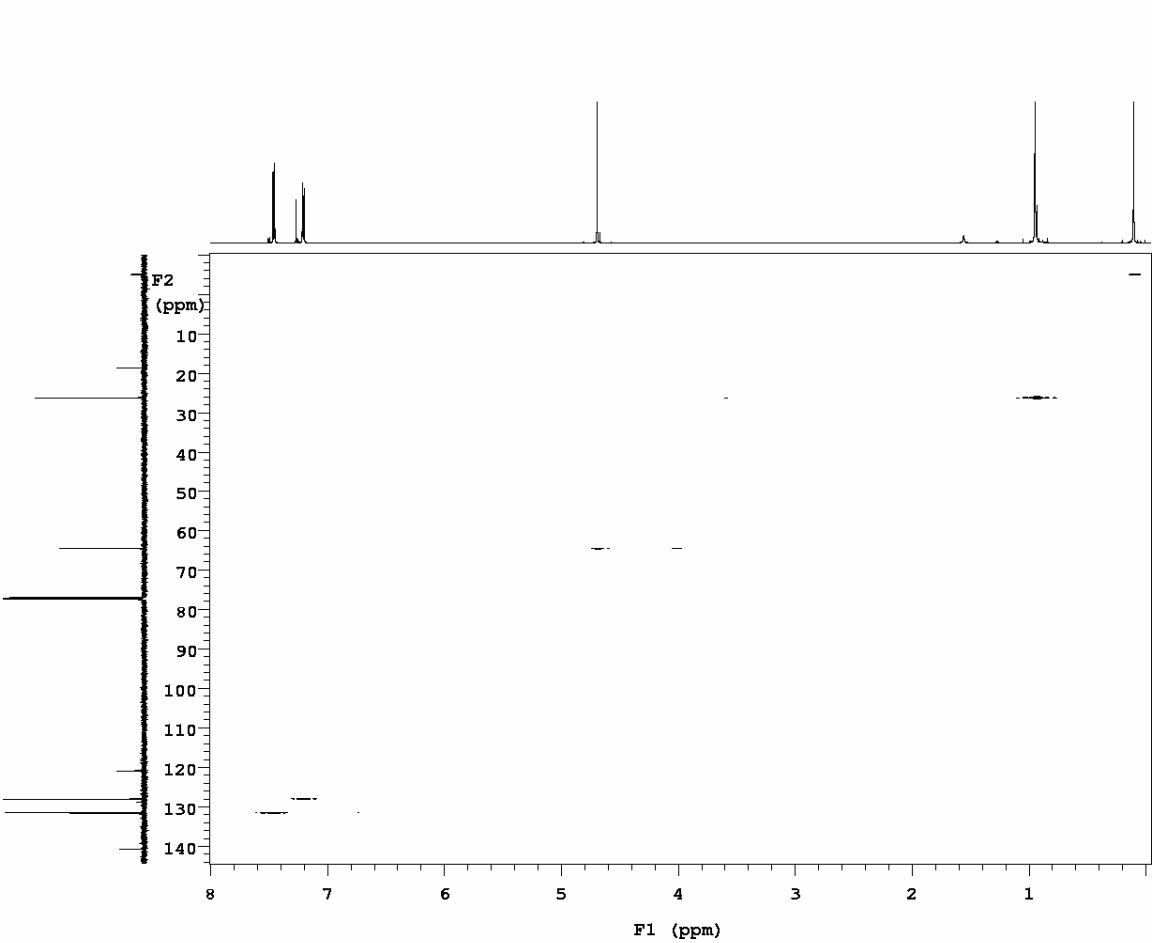
^{13}C NMR of 4-(TBS-O-methyl)-phenyl bromide (2)



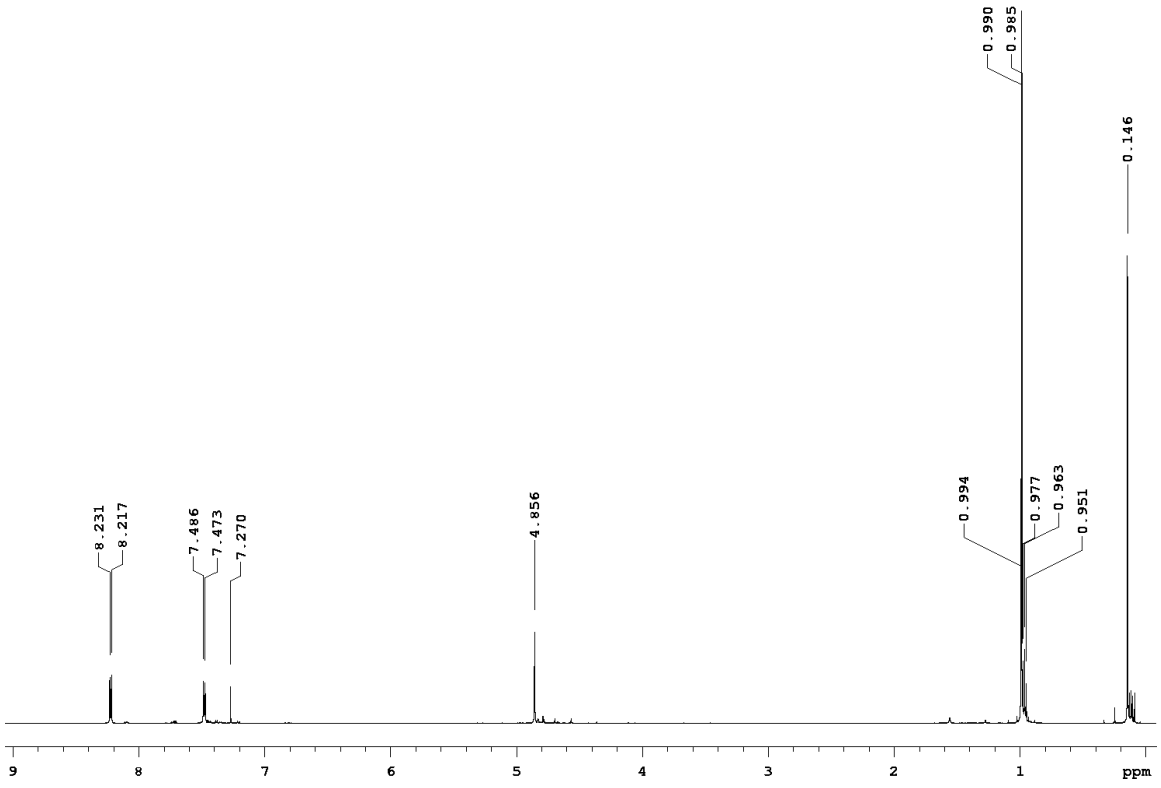
^{1}H-^{1}H COSY NMR of 4-(TBS-O-methyl)-phenyl bromide (2)



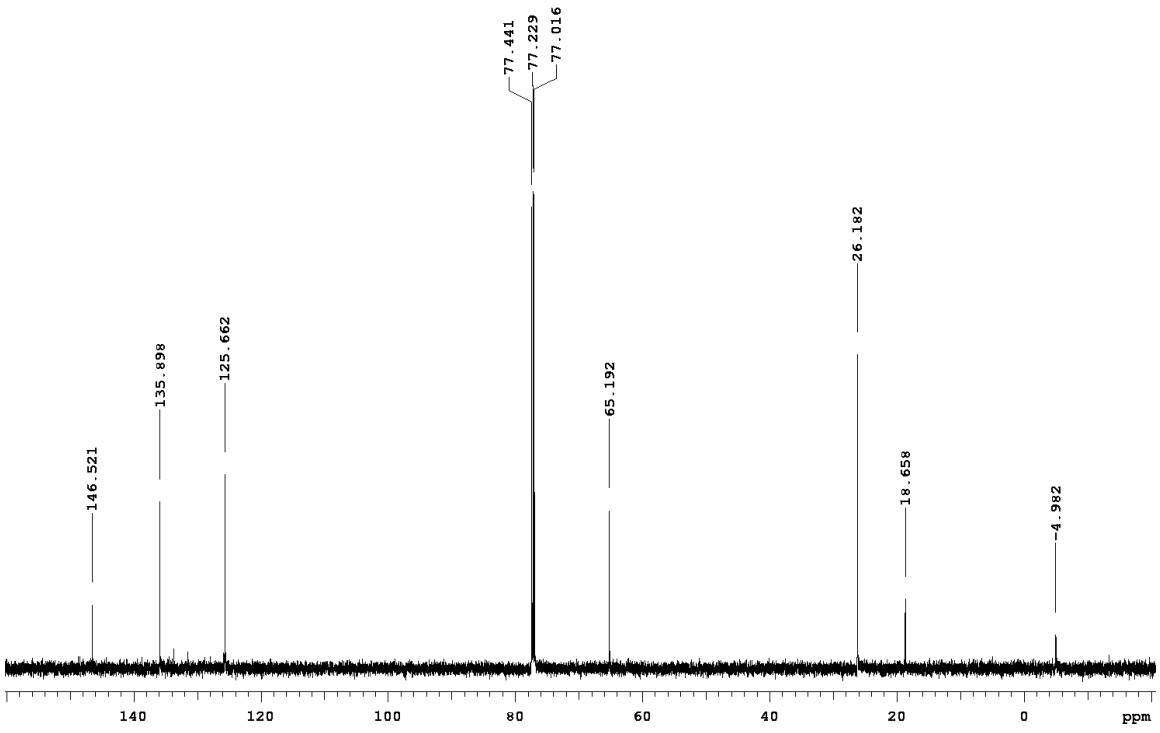
^{1}H-^{13}C HETCOR NMR of 4-(TBS-O-methyl)-phenyl bromide (2)



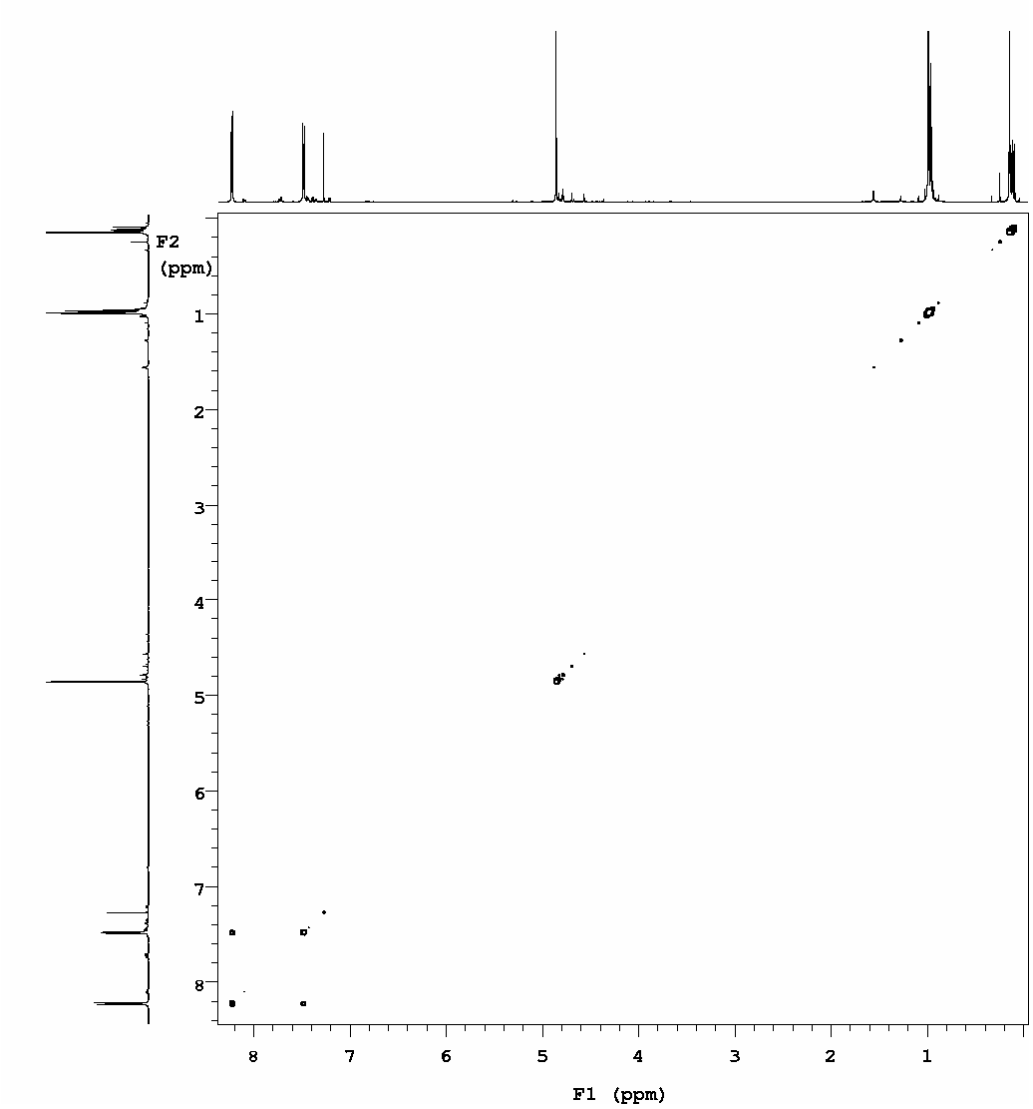
^{1}H NMR of 4-(TBS-O-methyl)-phenyl boronic acid (3)



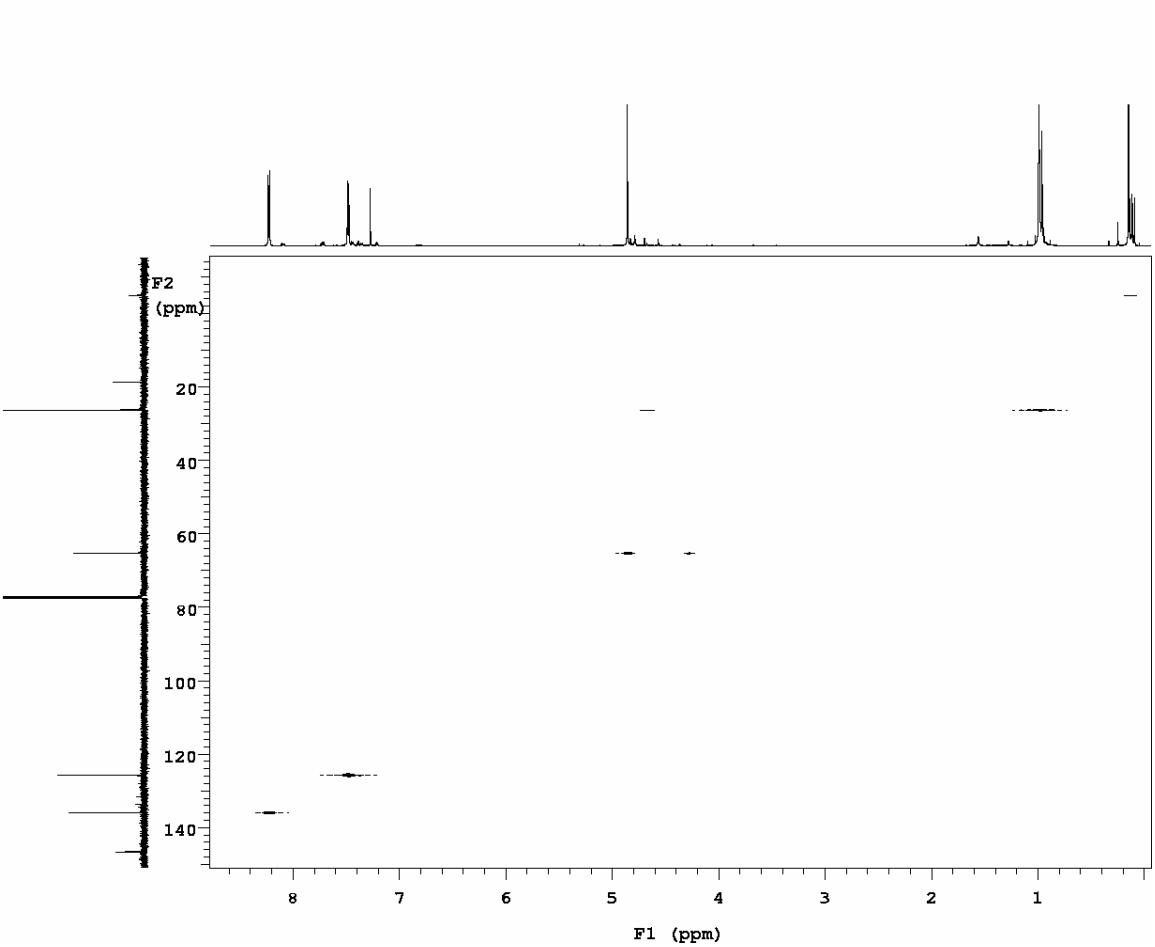
^{13}C NMR of 4-(TBS-O-methyl)-phenyl boronic acid (3)



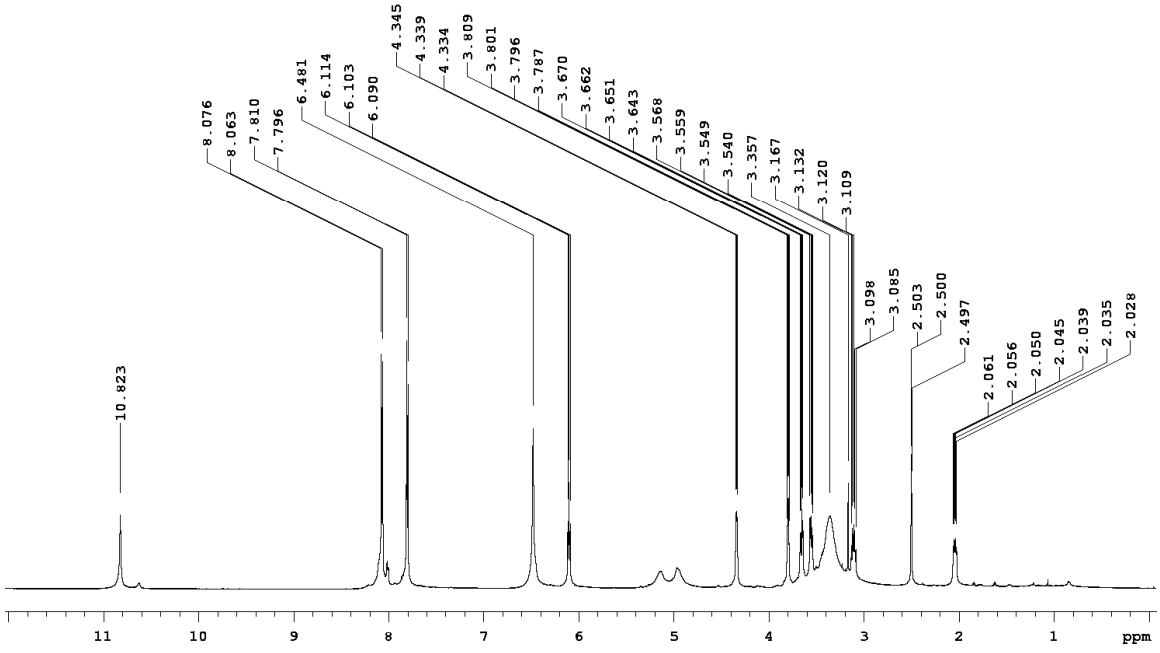
^{1}H-^{1}H COSY NMR of 4-(TBS-O-methyl)-phenyl boronic acid (3)



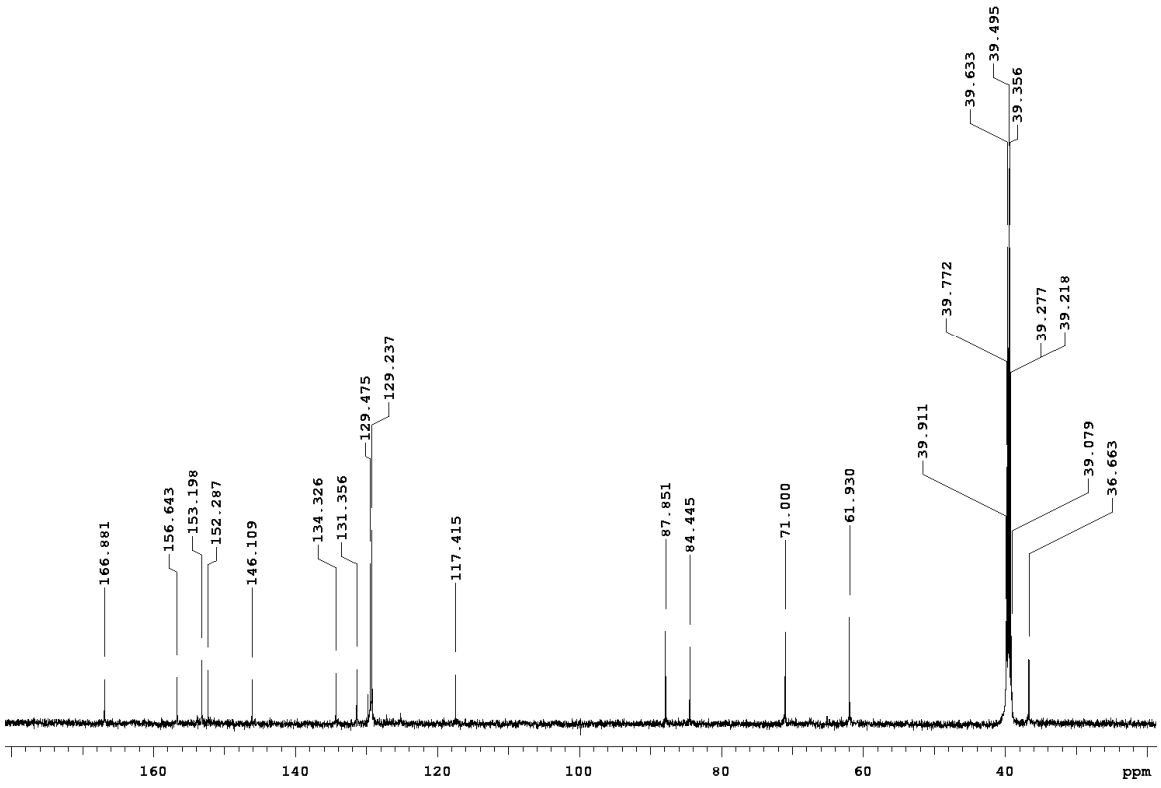
^{1}H-^{13}C HETCOR NMR of 4-(TBS-O-methyl)-phenyl boronic acid (3)



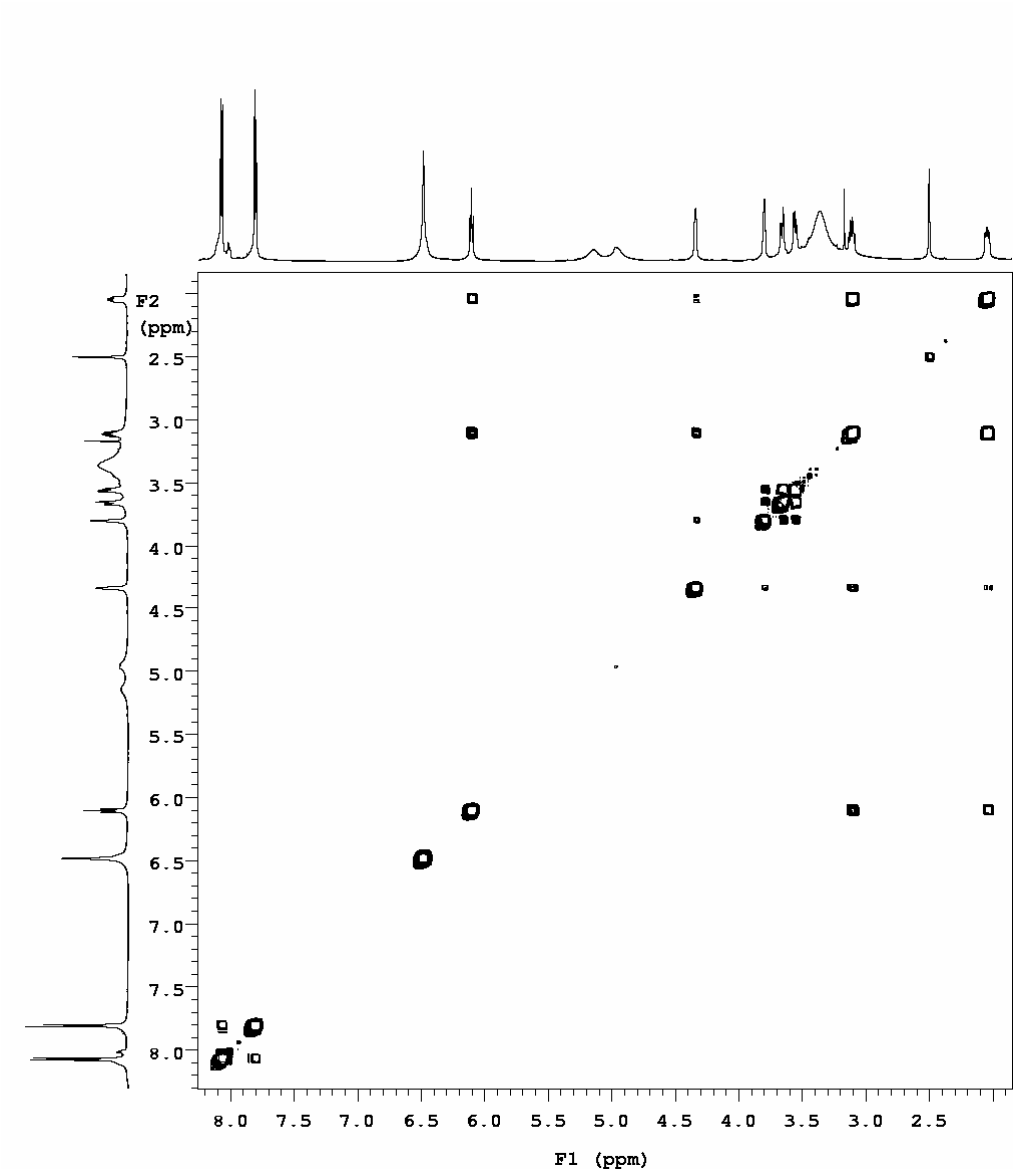
^{1}H NMR of 8-(4-Carboxyphenyl)-2′-deoxyguanosine (4)



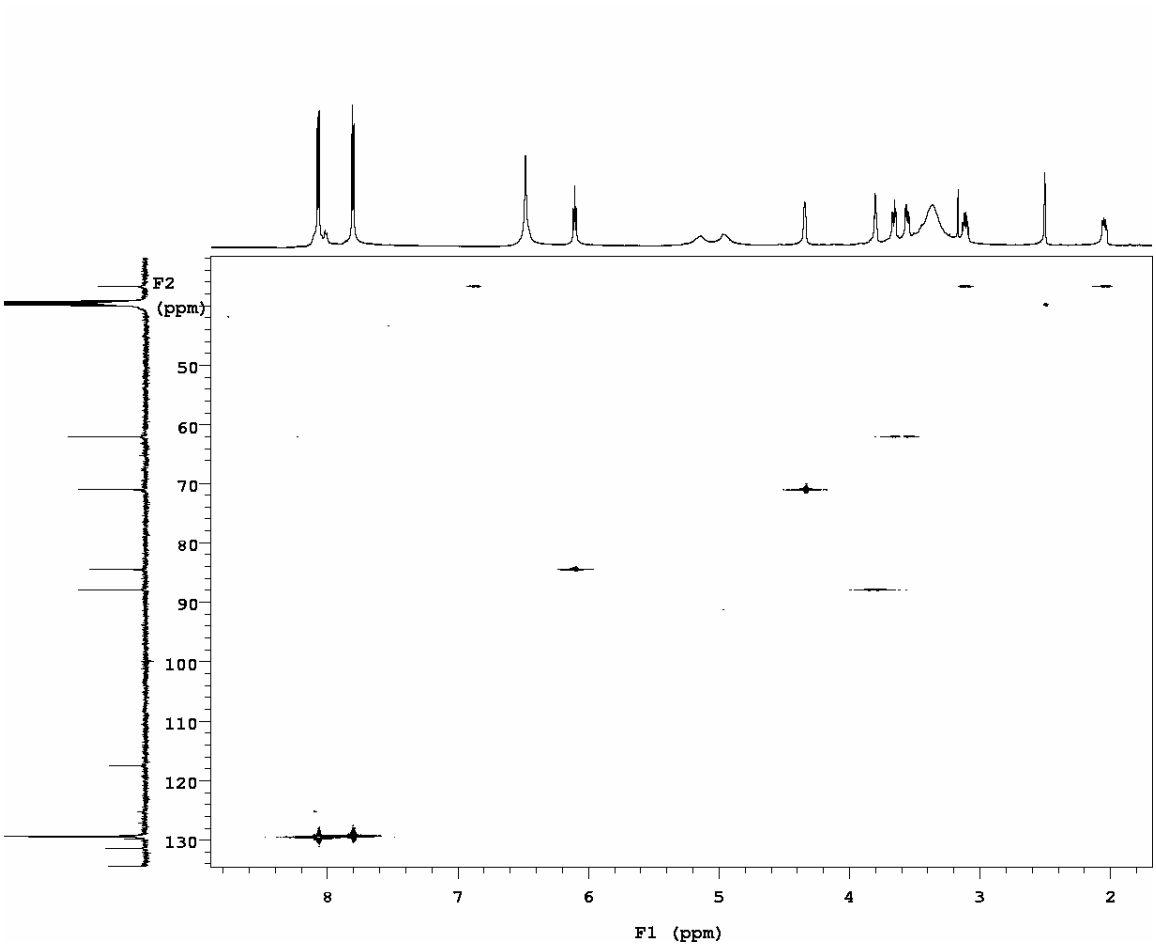
^{13}C NMR of 8-(4-Carboxyphenyl)-2′-deoxyguanosine (4)



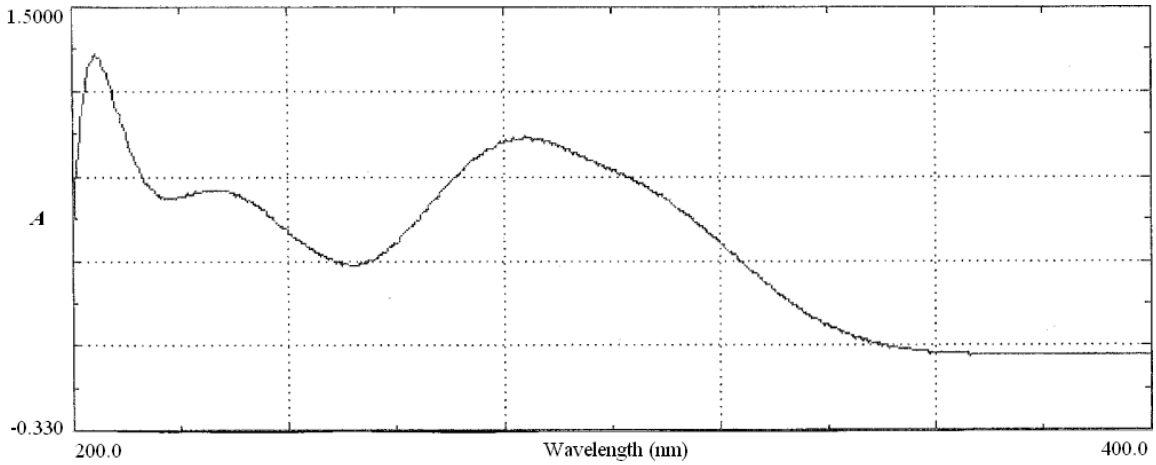
^{1}H-^{1}H COSY NMR of 8-(4-Carboxyphenyl)-2′-deoxyguanosine (4)



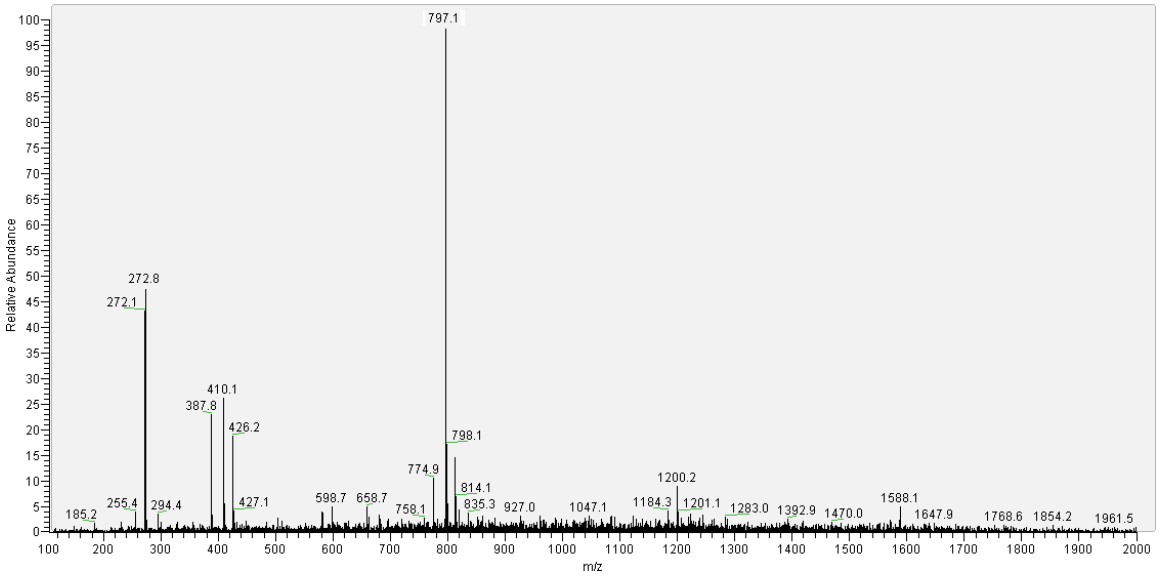
^{1}H-^{13}C HETCOR NMR of 8-(4-Carboxyphenyl)-2′-deoxyguanosine (4)



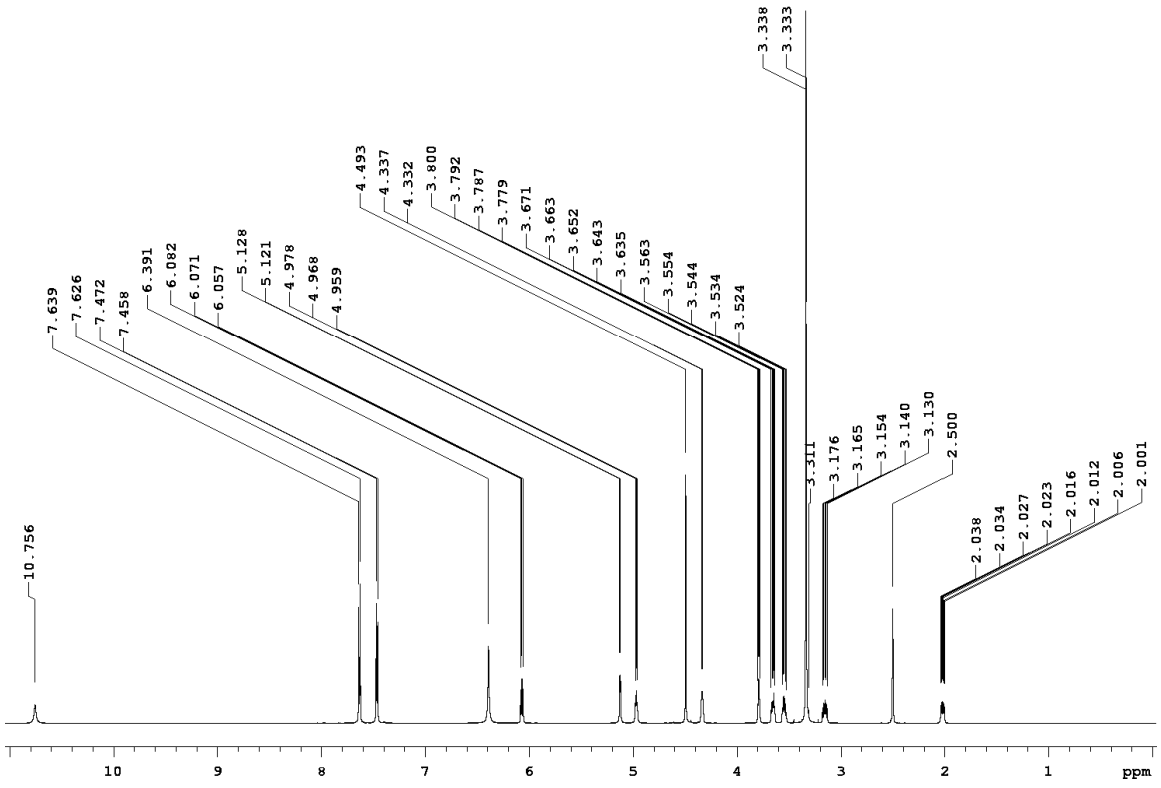
UV Spectrum of 8-(4-Carboxyphenyl)-2′-deoxyguanosine (4)



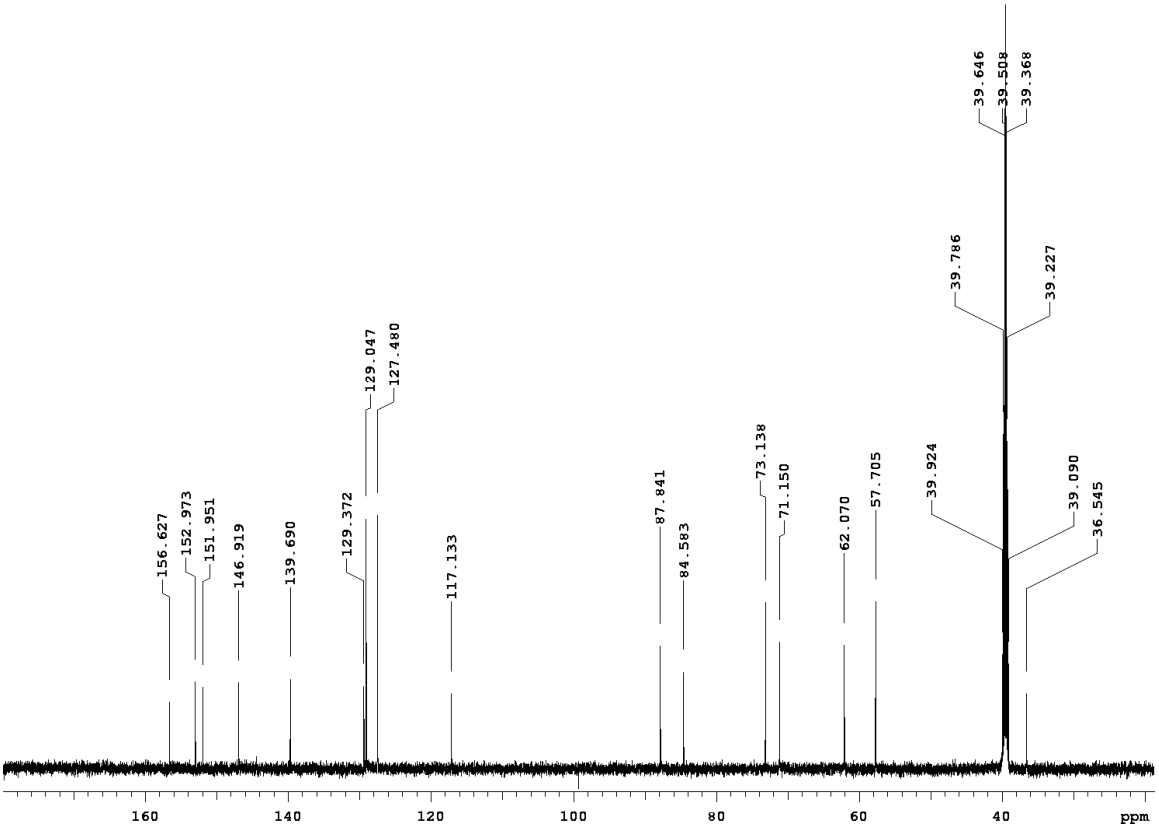
MS^+ of 8-(4-Carboxyphenyl)-2′-deoxyguanosine (4)



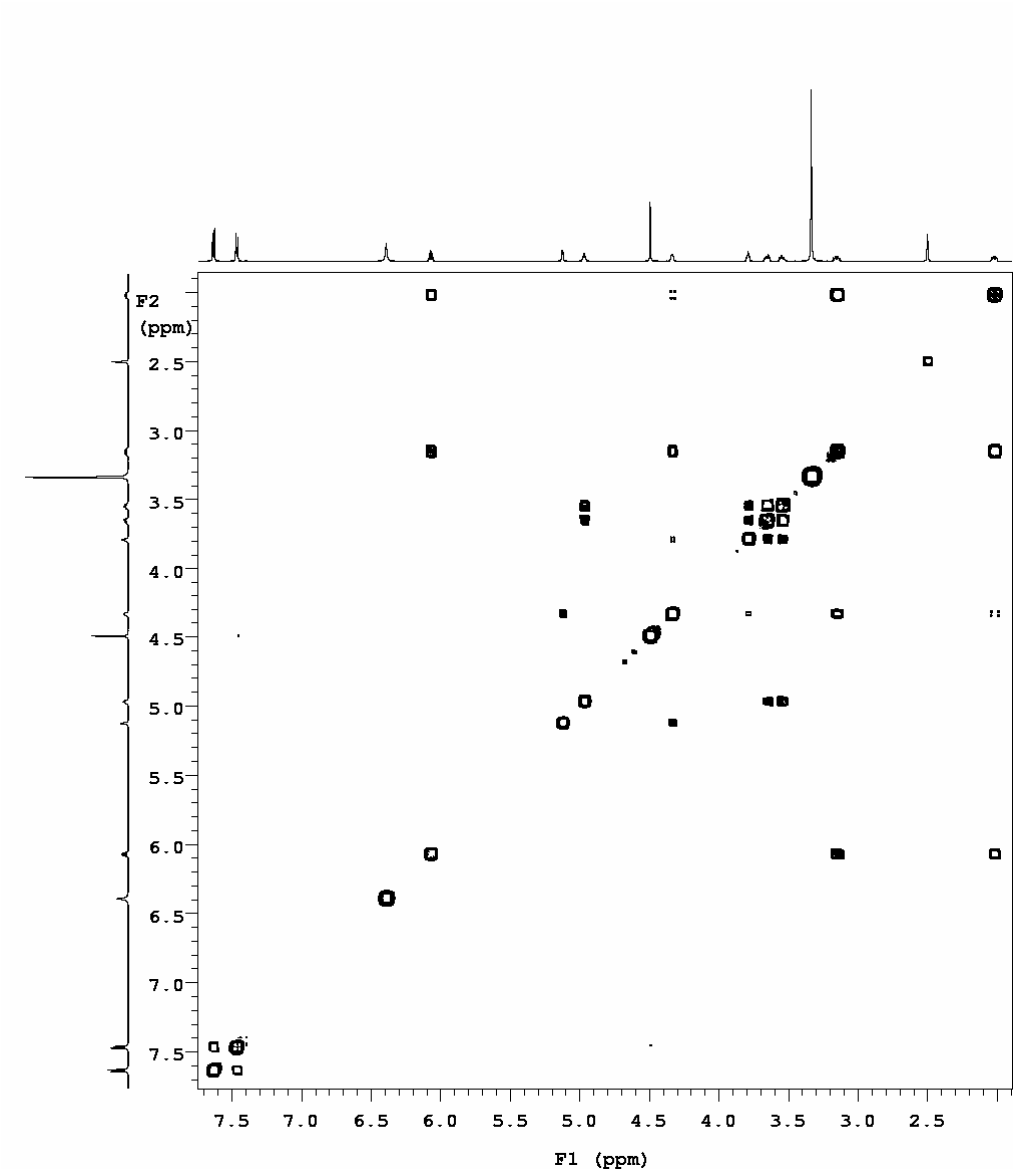
^{1}H NMR of 8-(4-Methoxymethylphenyl)-2′-deoxyguanosine (5)



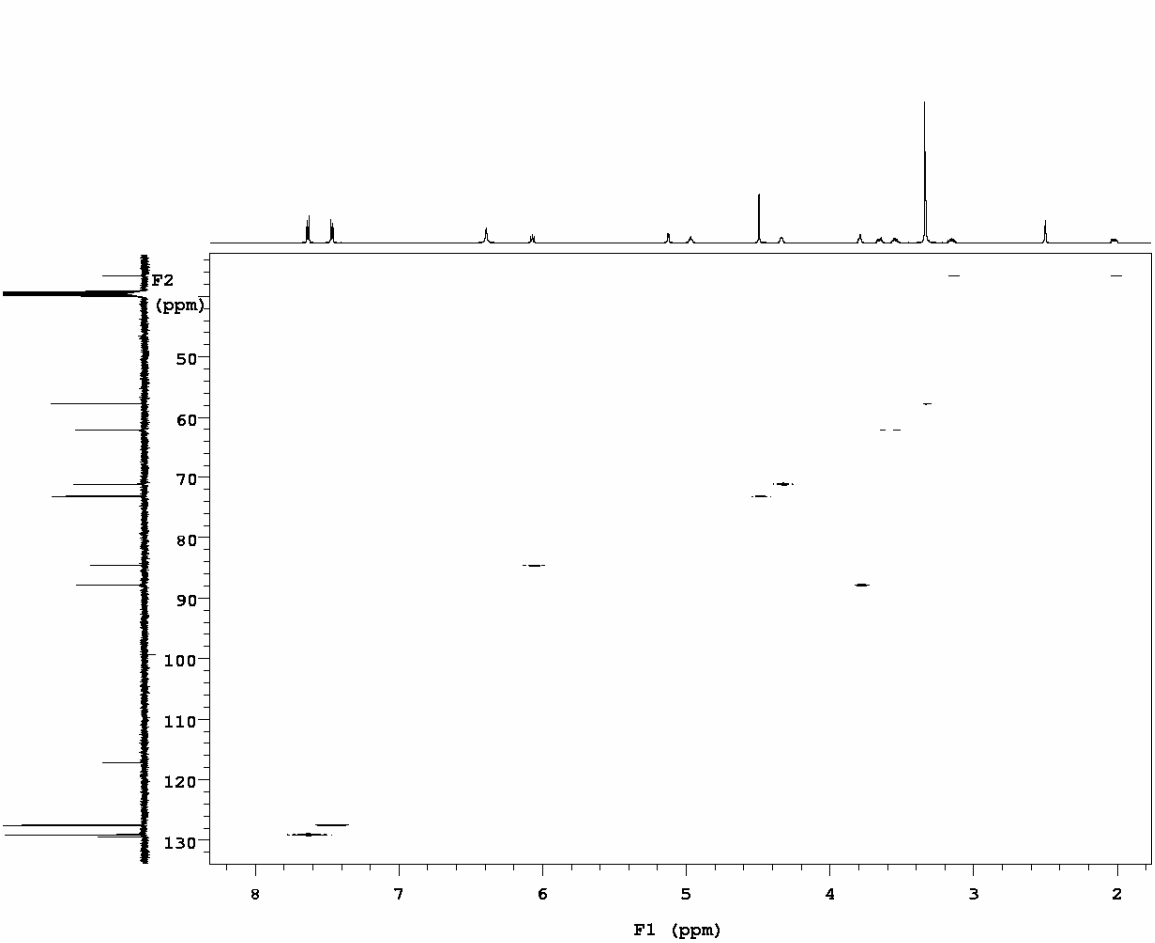
^{13}C NMR of 8-(4-Methoxymethylphenyl)-2′-deoxyguanosine (5)



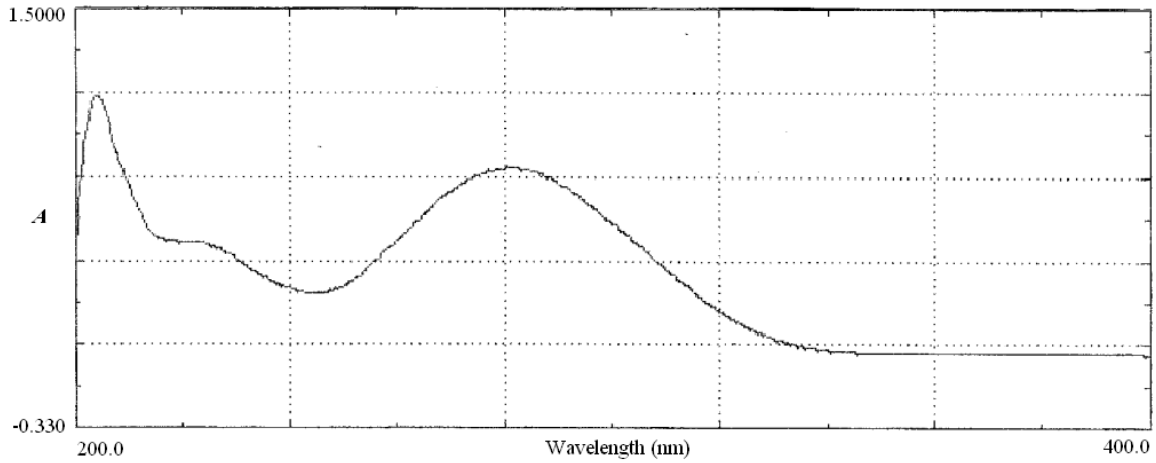
^{1}H-^{1}H COSY NMR of 8-(4-Methoxymethylphenyl)-2′-deoxyguanosine (5)



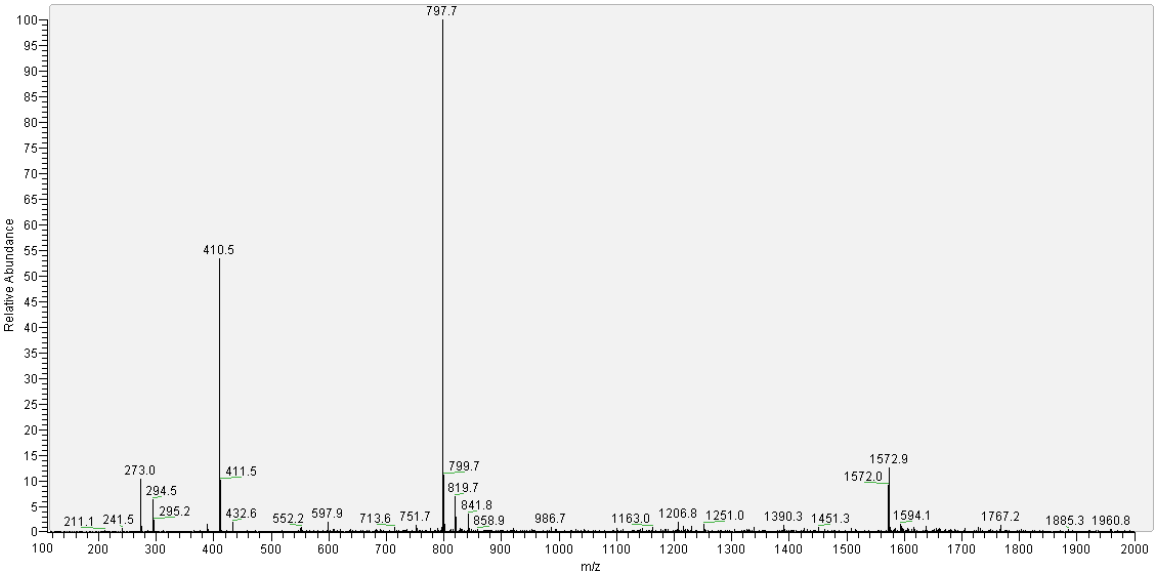
^{1}H-^{13}C HETCOR NMR of 8-(4-Methoxymethylphenyl)-2′-deoxyguanosine (5)



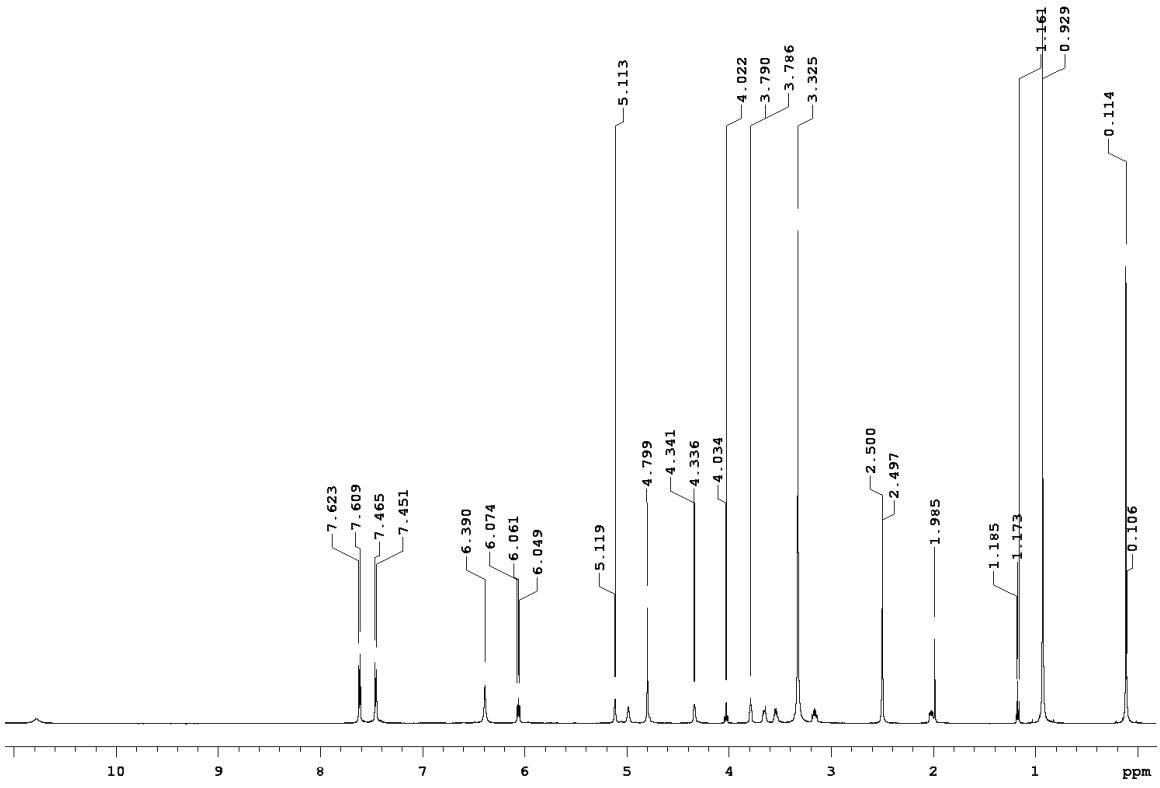
UV Spectrum of 8-(4-Methoxymethylphenyl)-2′-deoxyguanosine (5)



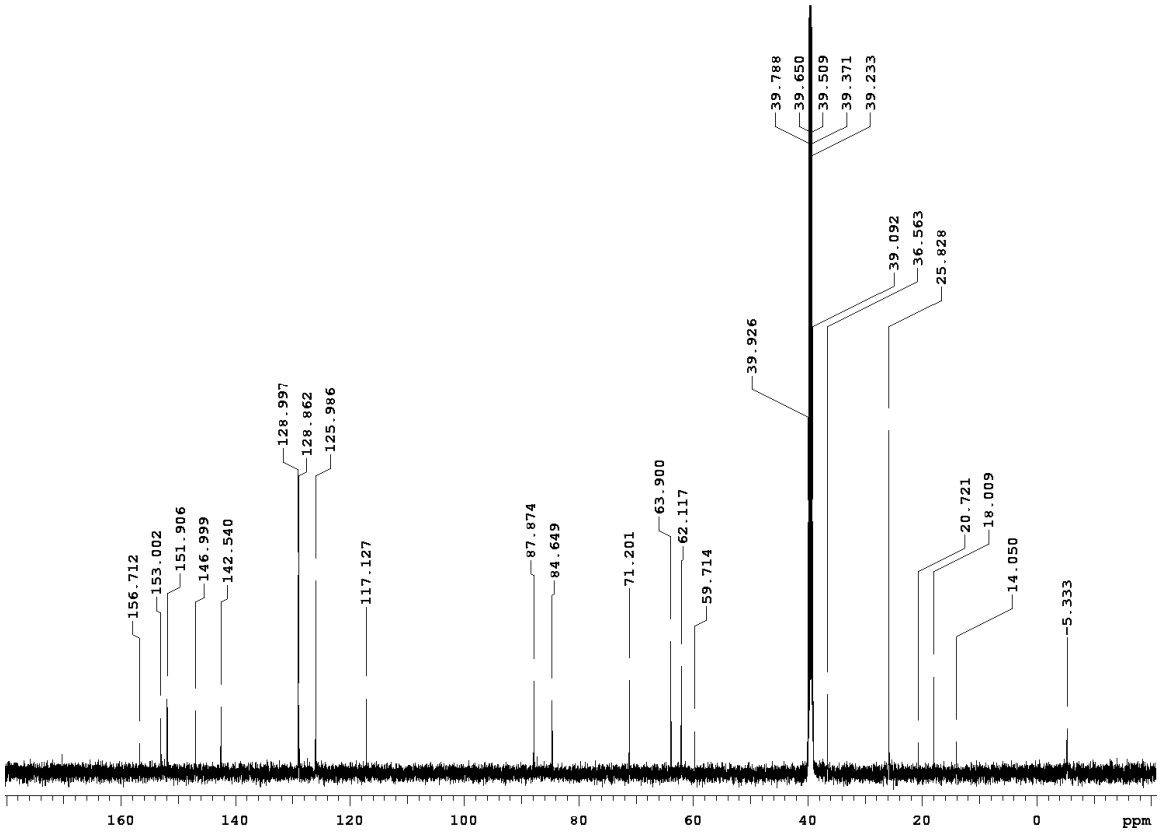
MS^+ of 8-(4-Methoxymethylphenyl)-2′-deoxyguanosine (5)



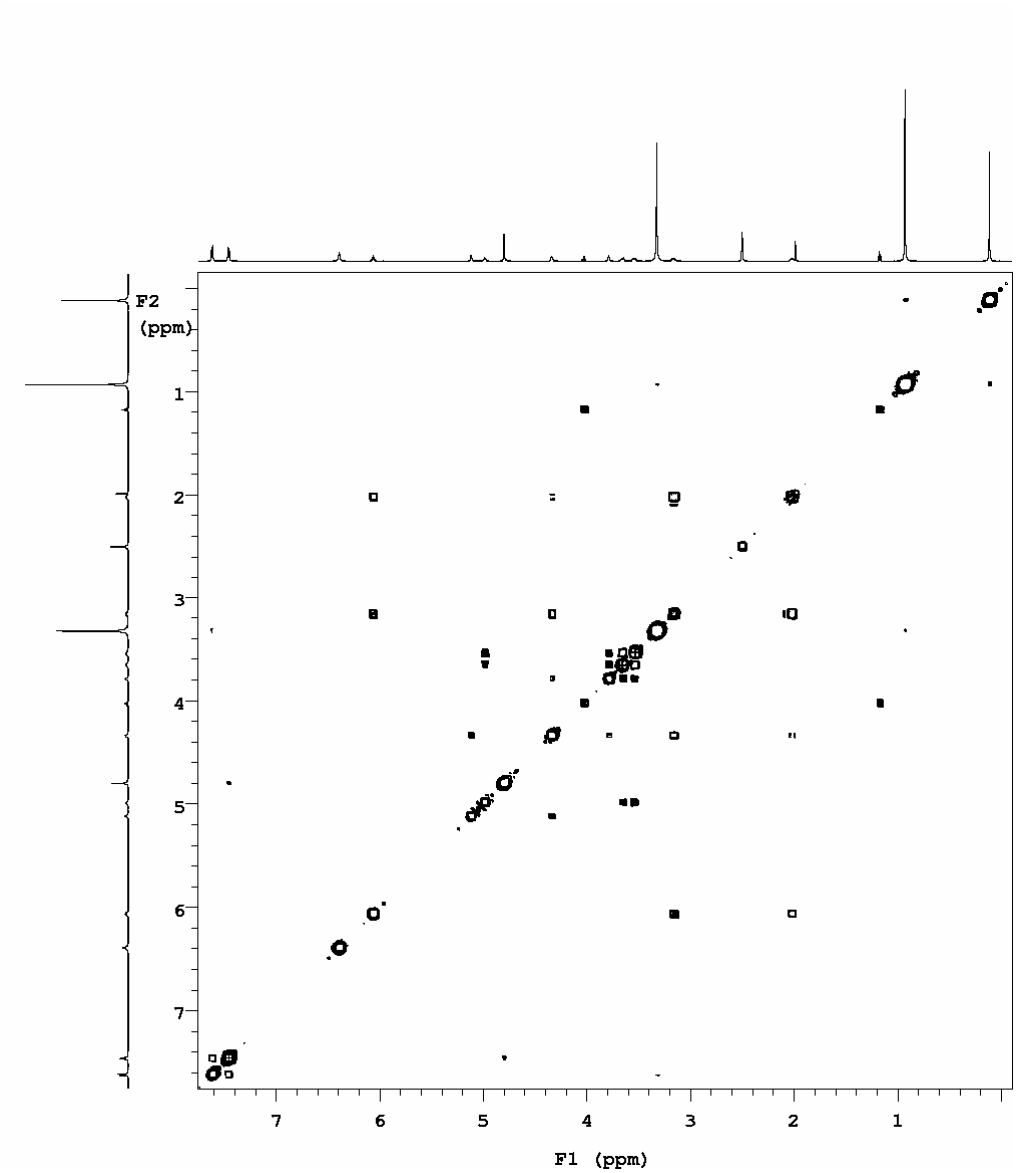
^{1}H NMR of 8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine (6)



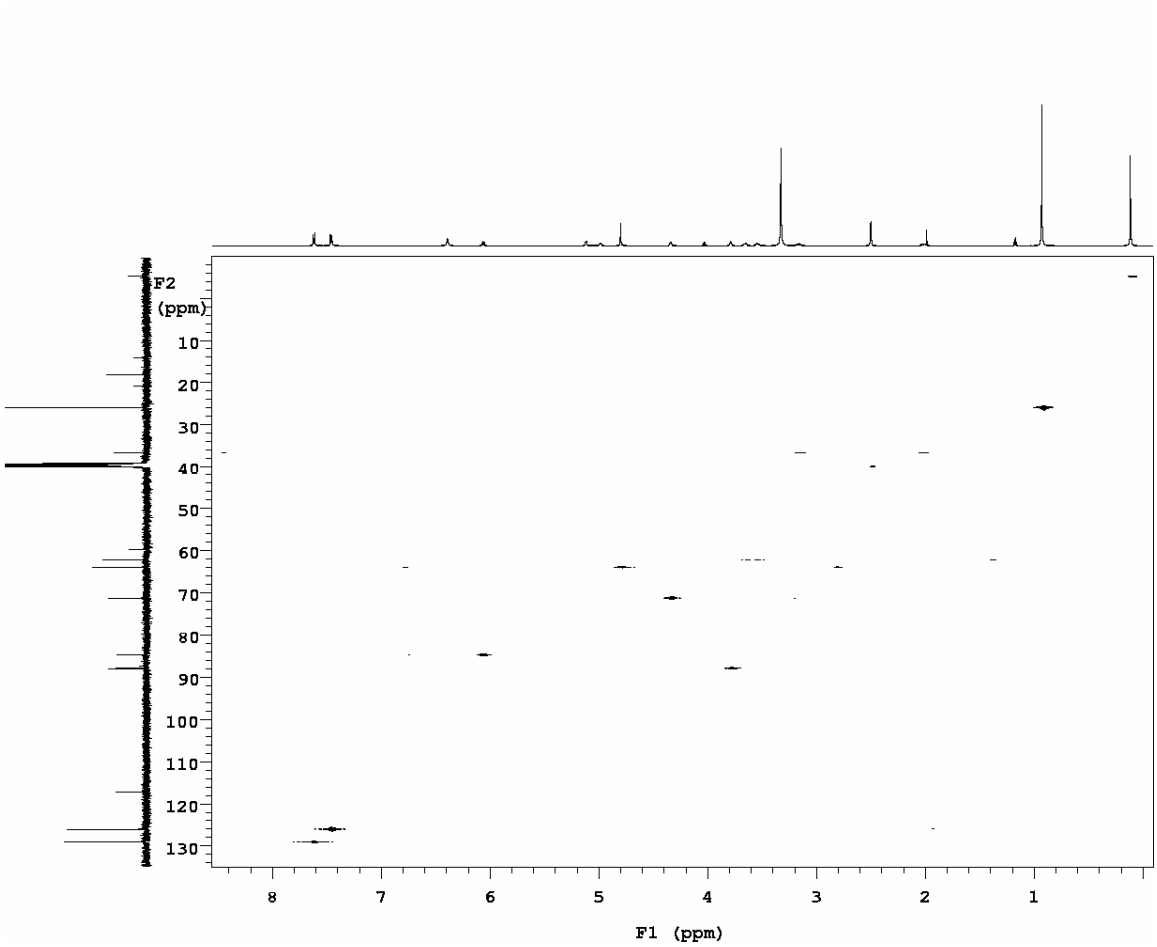
^{13}C NMR of 8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine (6)



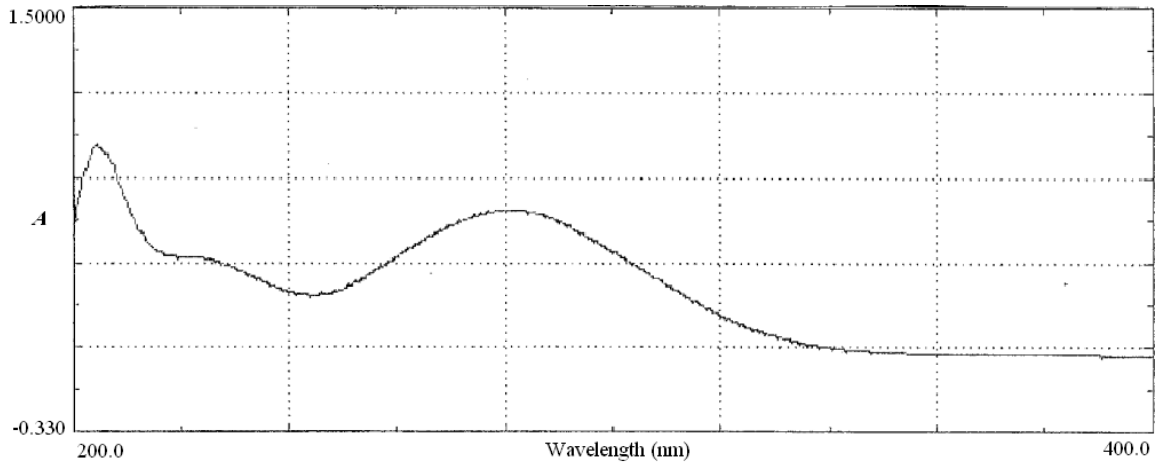
^{1}H-^{1}H COSY NMR of 8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine (6)



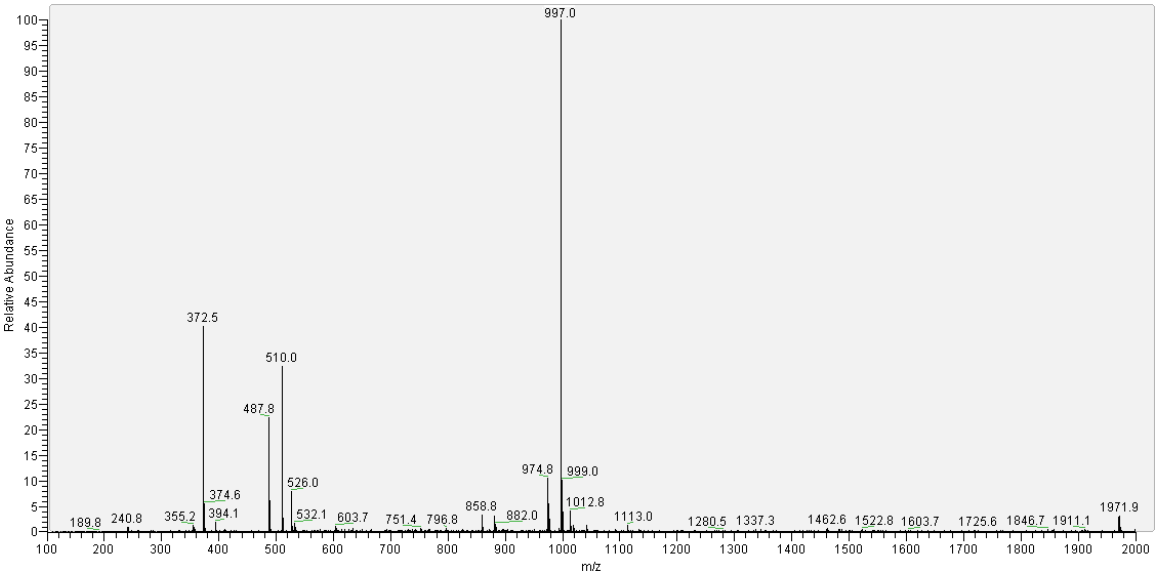
^{1}H-^{13}C HETCOR NMR of 8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine (6)



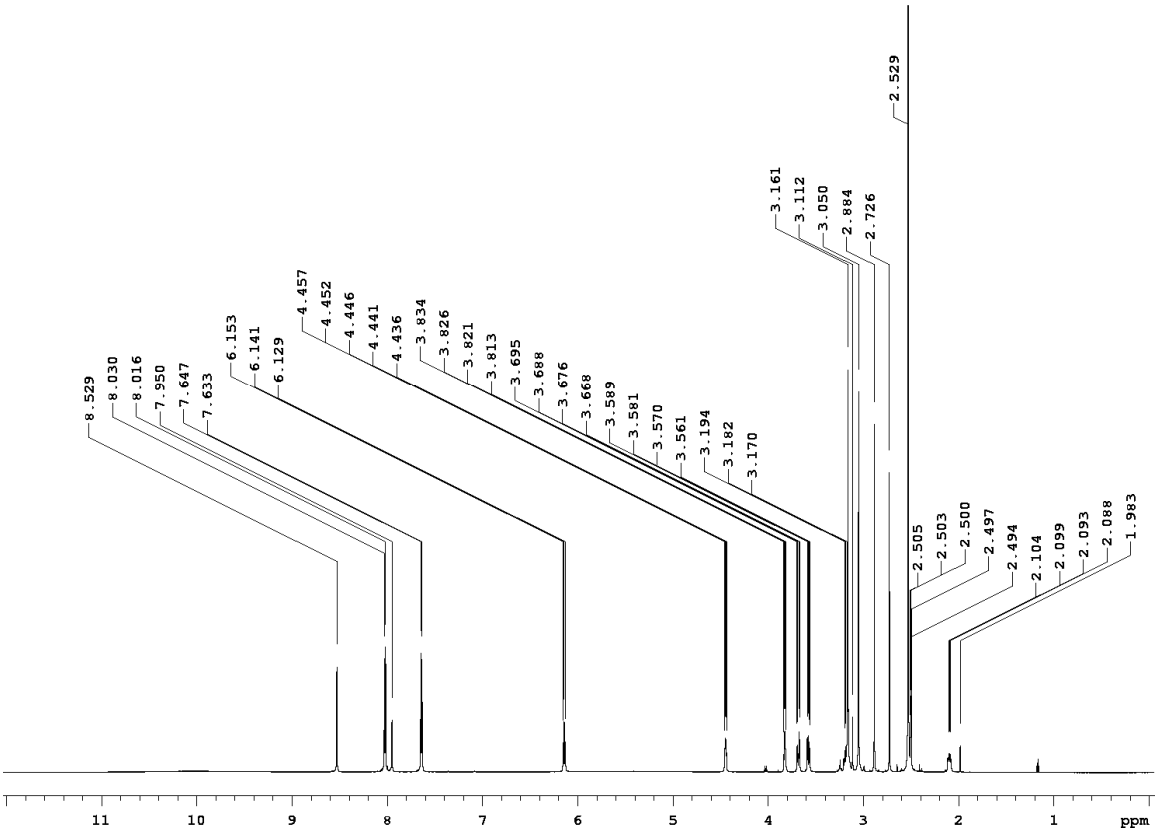
UV Spectrum of 8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine (6)



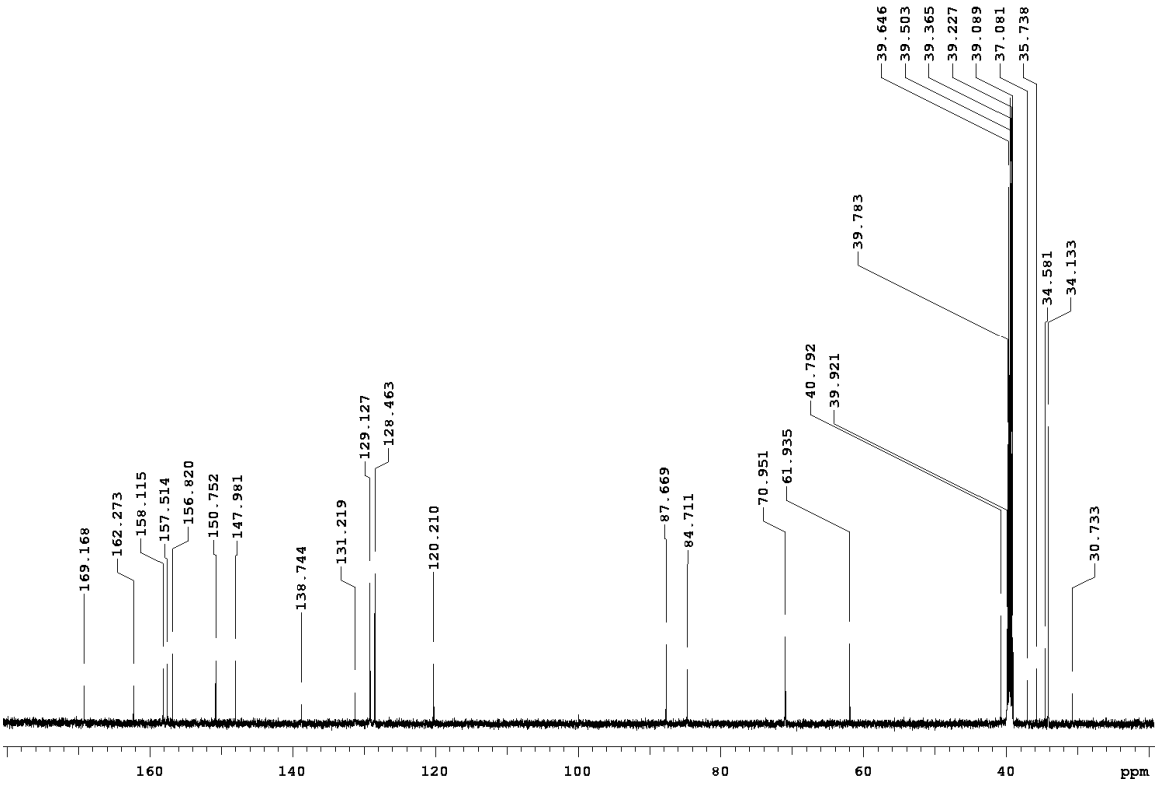
MS$^+$ of 8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine (6)



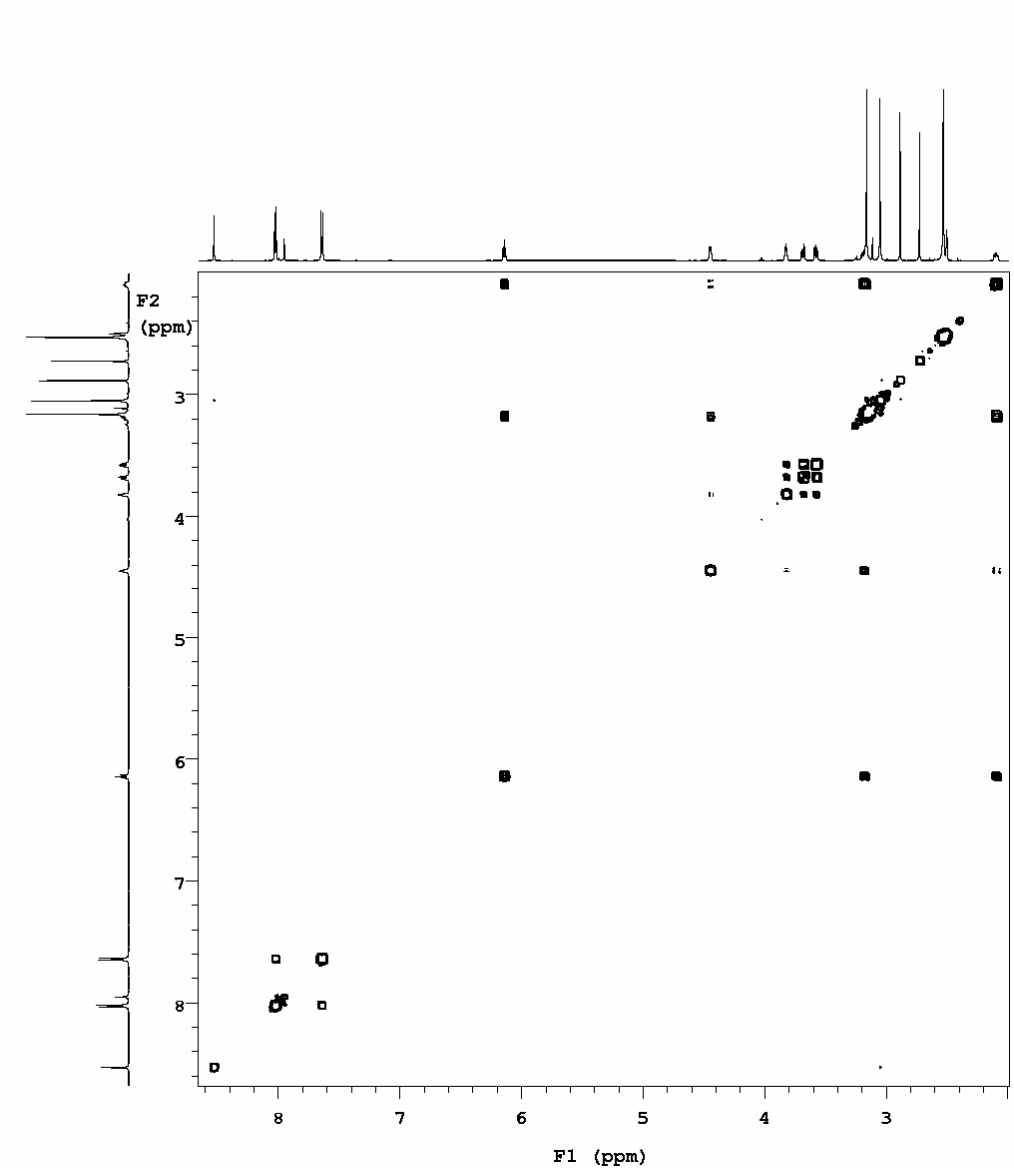
^{1}H NMR of N2-(N,N-Dimethylformamidine)-8-(4-carboxyphenyl)-2′-deoxyguanosine (7)



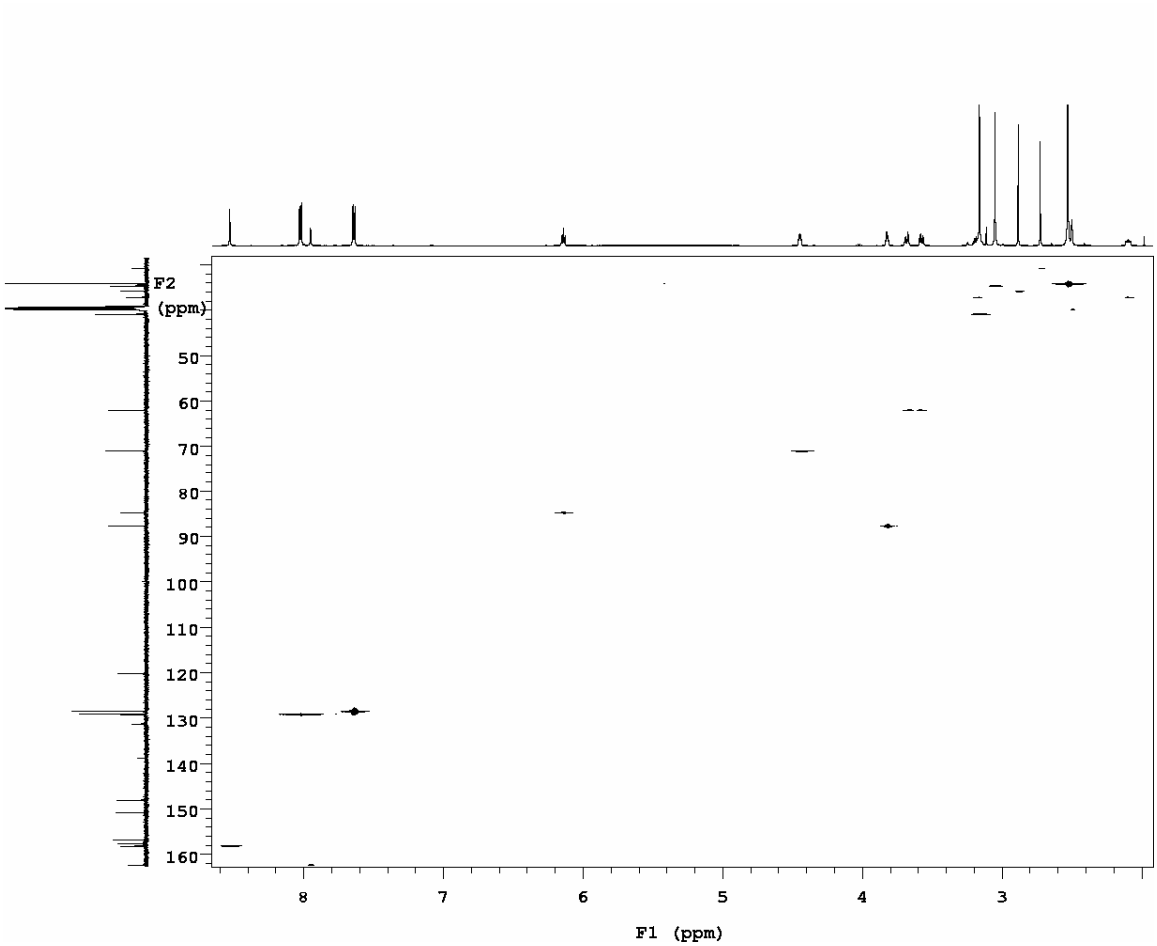
^{13}C NMR of N2-(N,N-Dimethylformamidine)-8-(4-carboxyphenyl)-2′-deoxyguanosine (7)



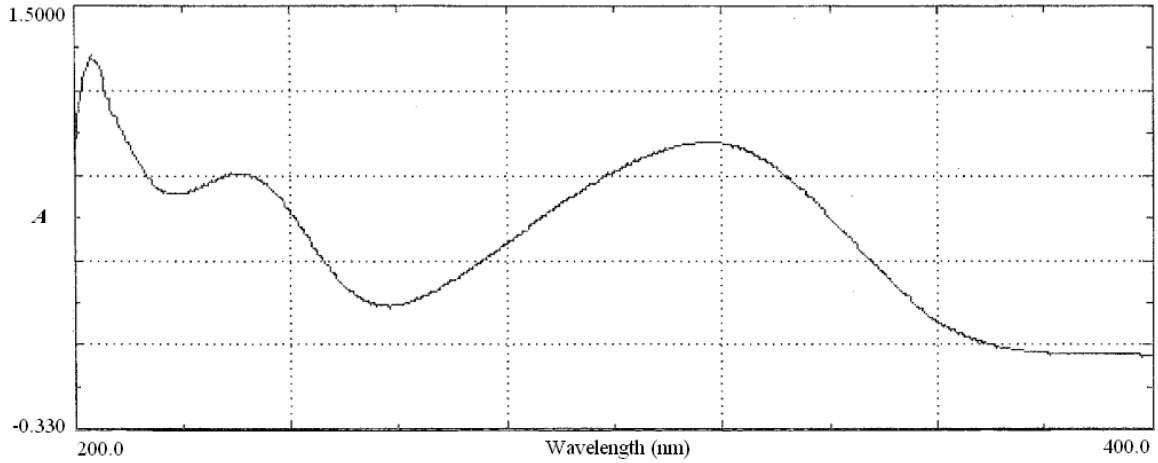
^{1}H-^{1}H COSY NMR of N2-(N,N-Dimethylformamidine)-8-(4-carboxyphenyl)-2′-deoxyguanosine (7)



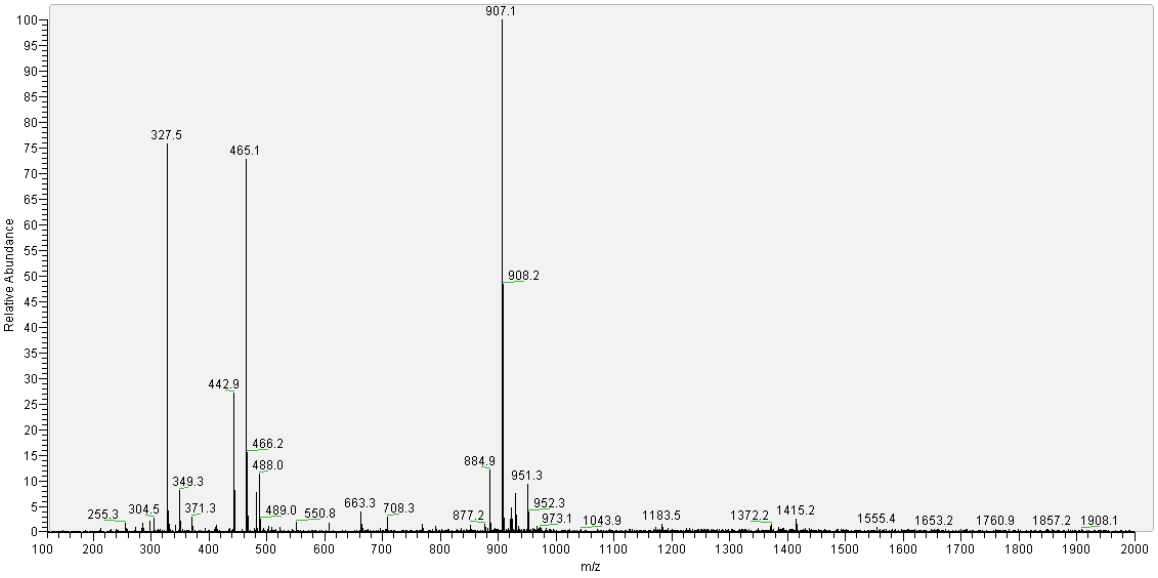
^{1}H-^{13}C HETCOR NMR of N2-(N,N-Dimethylformamidine)-8-(4-carboxyphenyl)-2′-deoxyguanosine (7)



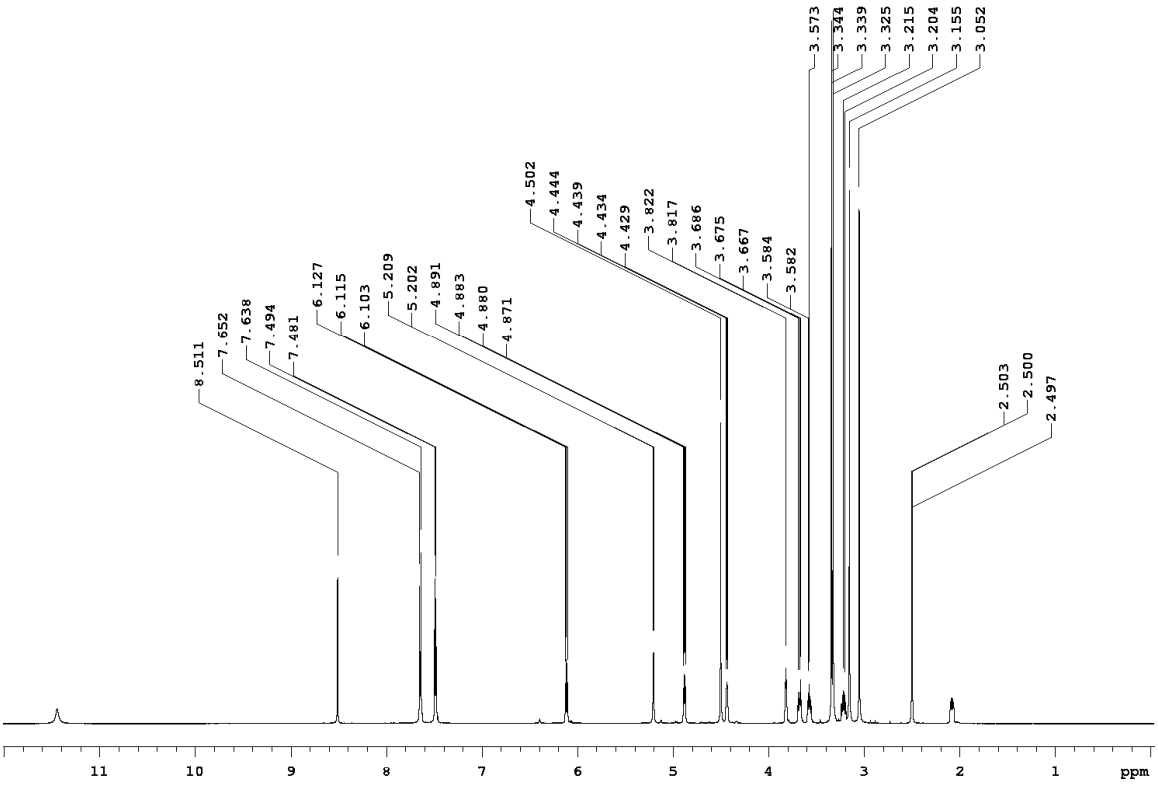
UV Spectrum of N2-(N,N-Dimethylformamidine)-8-(4-carboxyphenyl)-2′-deoxyguanosine (7)



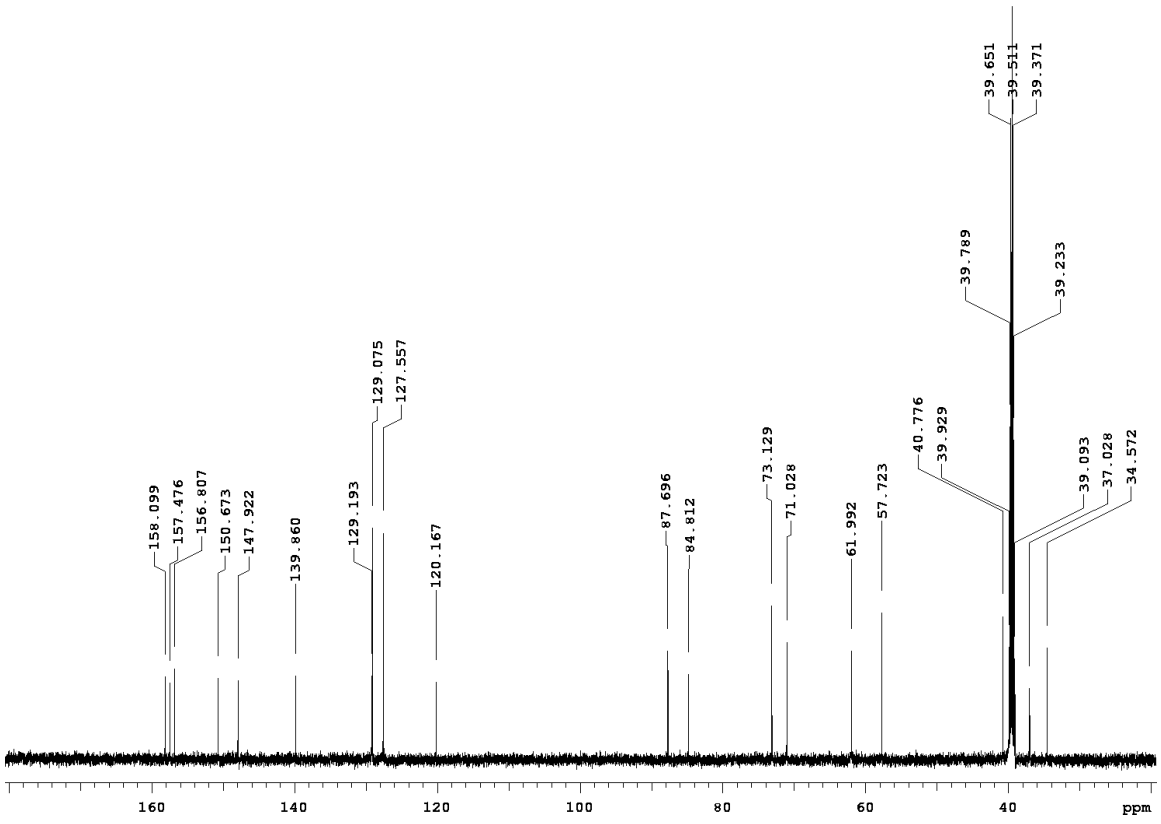
MS$^+$ of N2-(N,N-Dimethylformamidine)-8-(4-carboxyphenyl)-2′-deoxyguanosine (7)



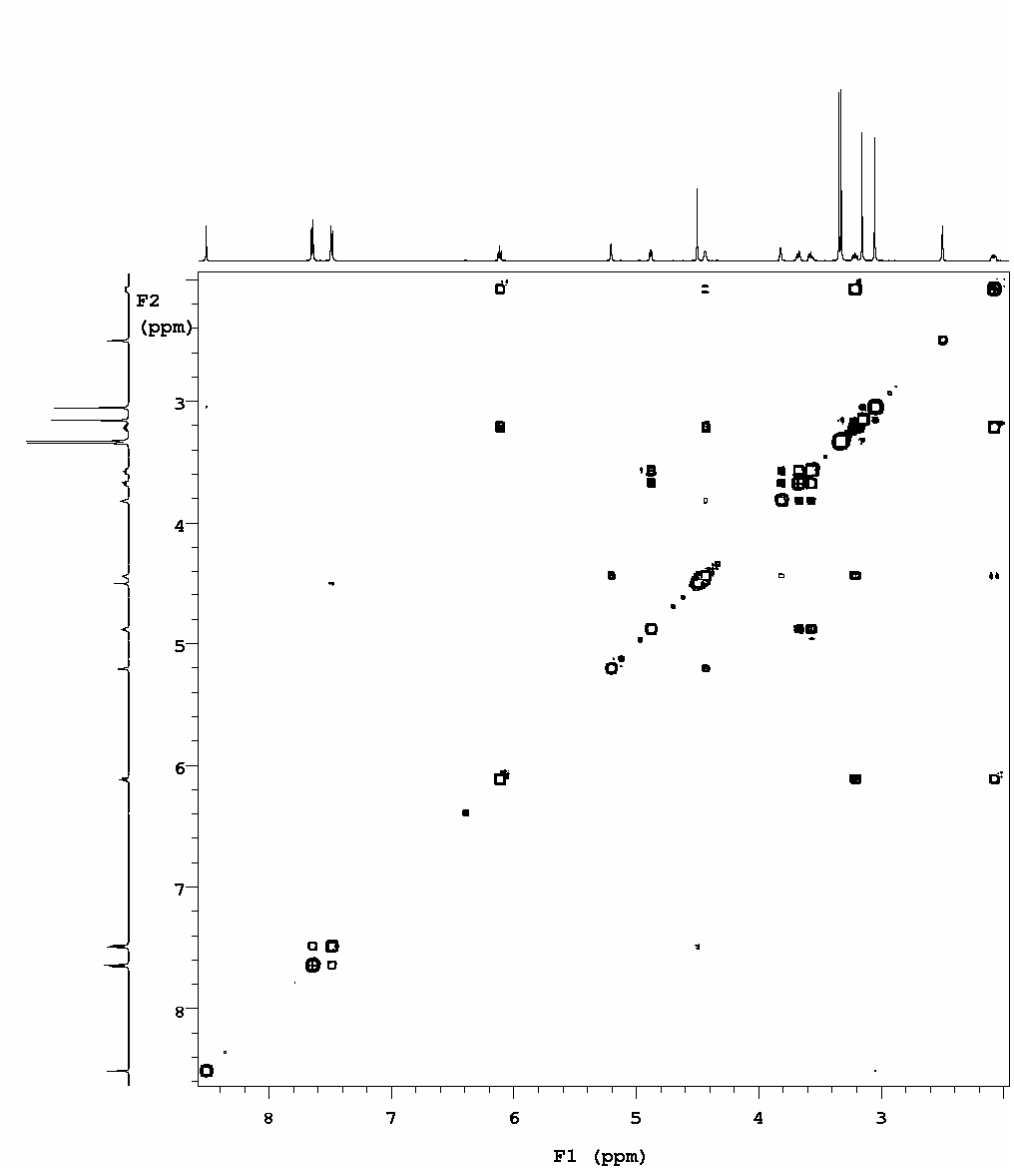
^{1}H NMR of N2-(N,N-Dimethylformamidine)-8-(4-methoxymethylphenyl)-2′-deoxyguanosine (8)



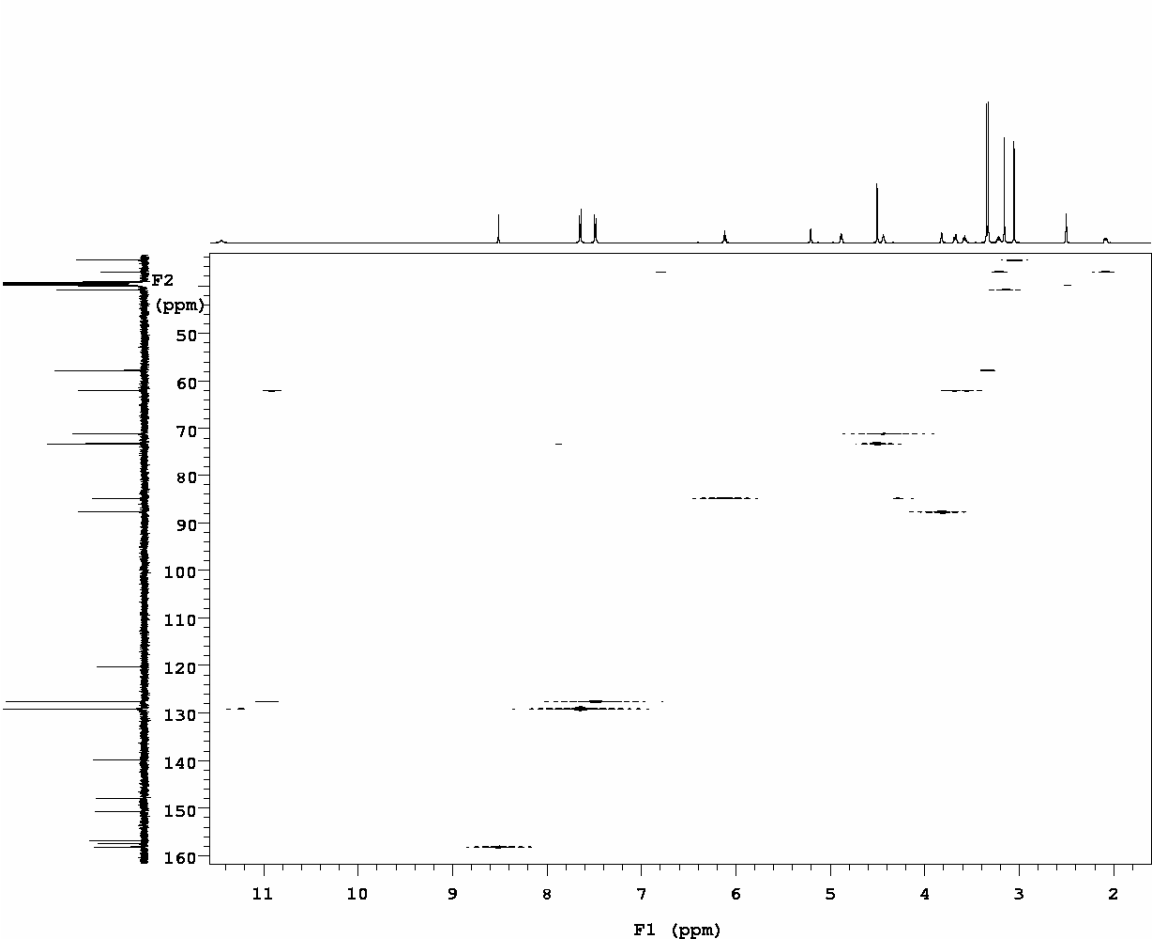
^{13}C NMR of N2-(N,N-Dimethylformamidine)-8-(4-methoxymethylphenyl)-2′-deoxyguanosine (8)



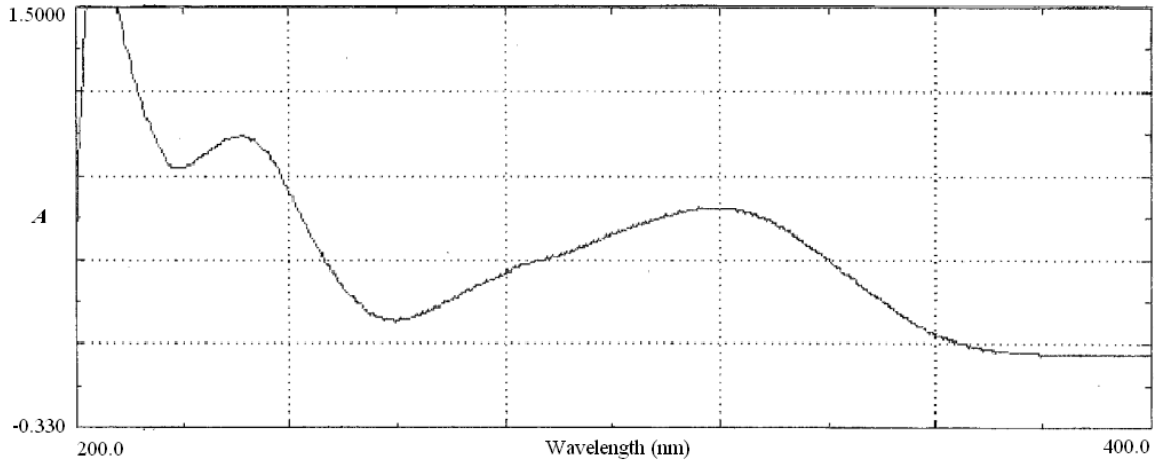
^{1}H-^{1}H COSY NMR of N2-(N,N-Dimethylformamidine)-8-(4-methoxymethylphenyl)-2′-deoxyguanosine (8)



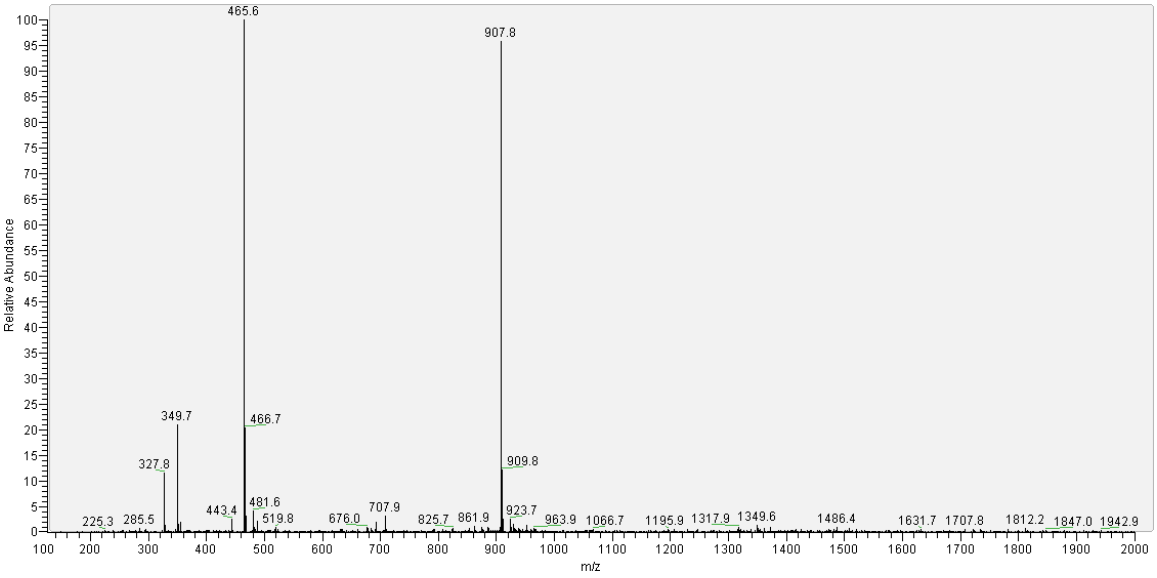
^{1}H-^{13}C HETCOR NMR of N2-(N,N-Dimethylformamidine)-8-(4-methoxymethylphenyl)-2′-deoxyguanosine (8)



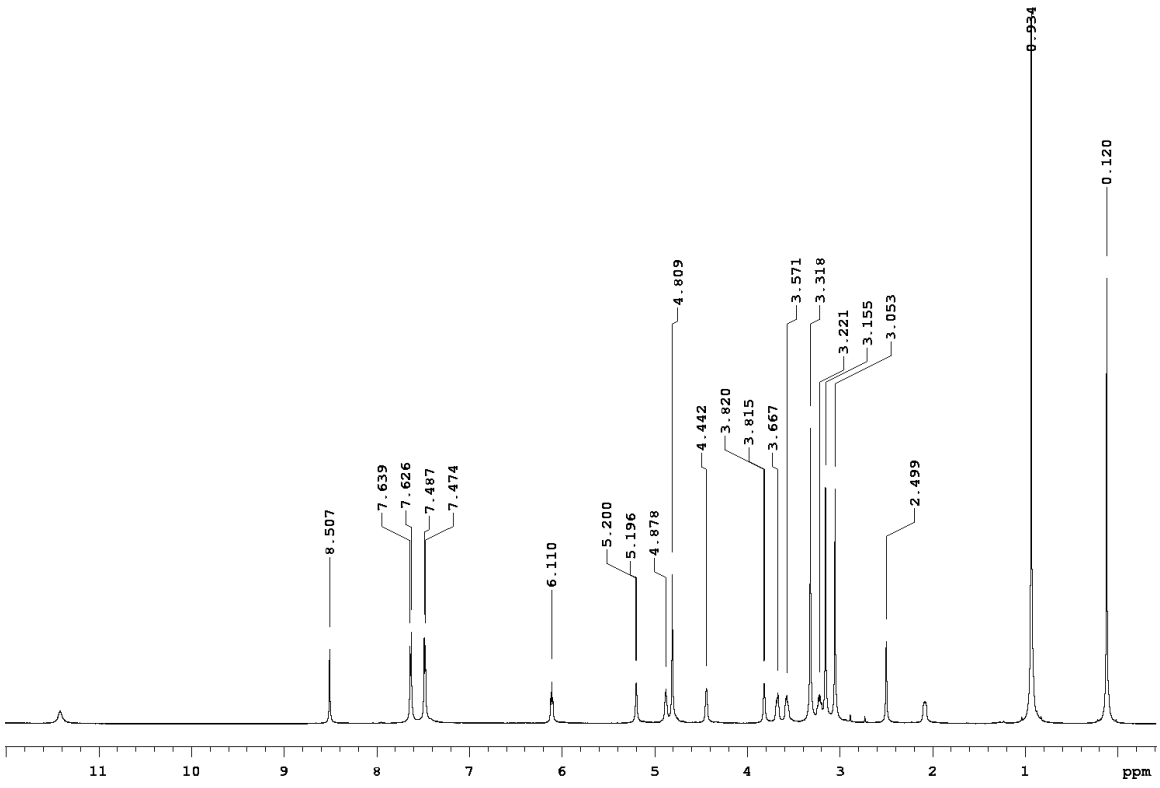
UV Spectrum of N2-(N,N-Dimethylformamidine)-8-(4-methoxymethylphenyl)-2′-deoxyguanosine (8)



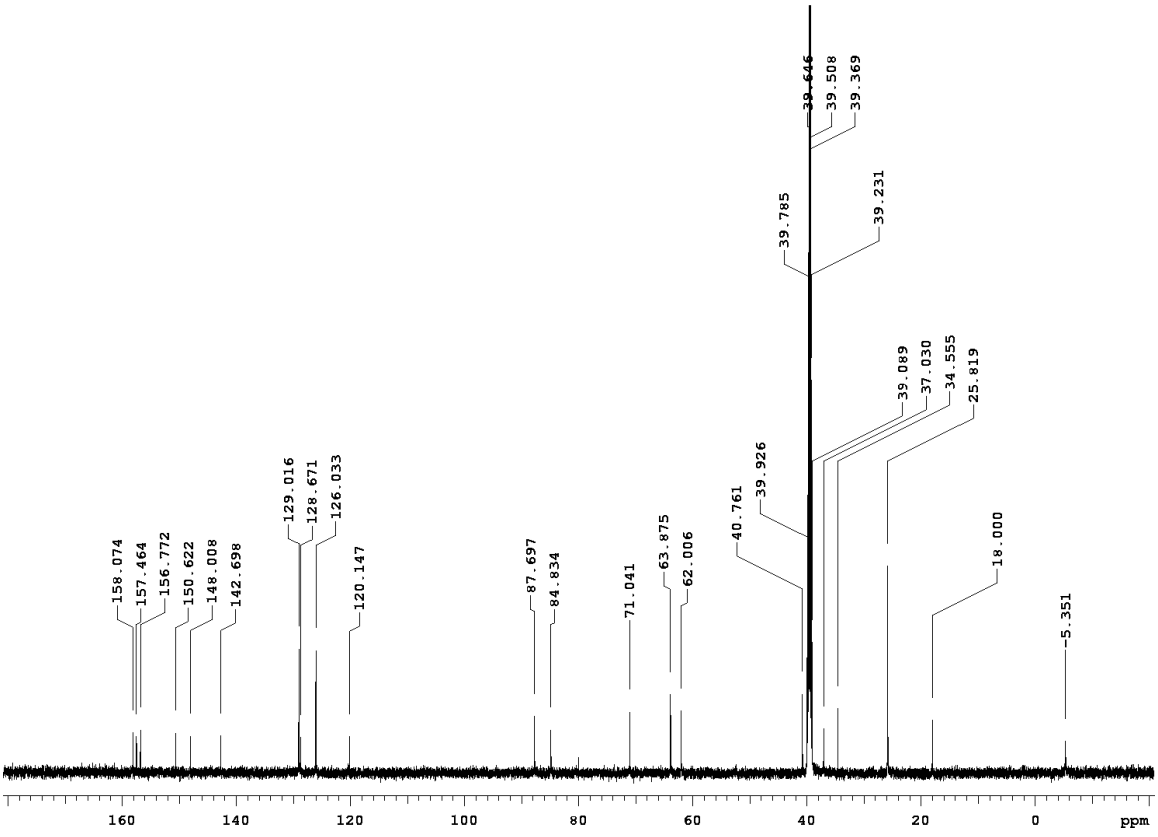
MS$^+$ of N2-(N,N-Dimethylformamidine)-8-(4-methoxymethylphenyl)-2′-deoxyguanosine (8)



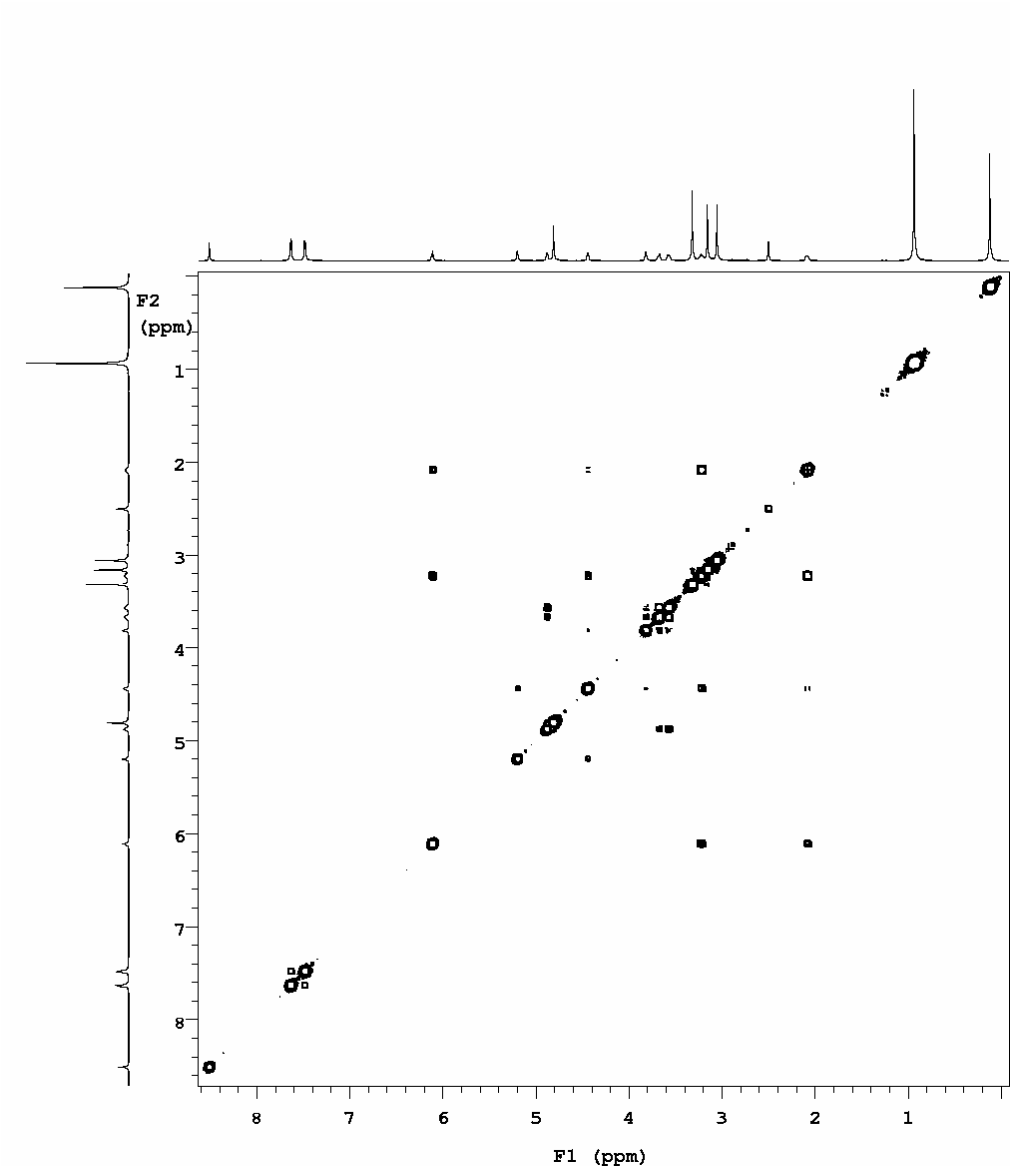
^{1}H NMR of N2-(N,N-Dimethylformamidine)-8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine (9)



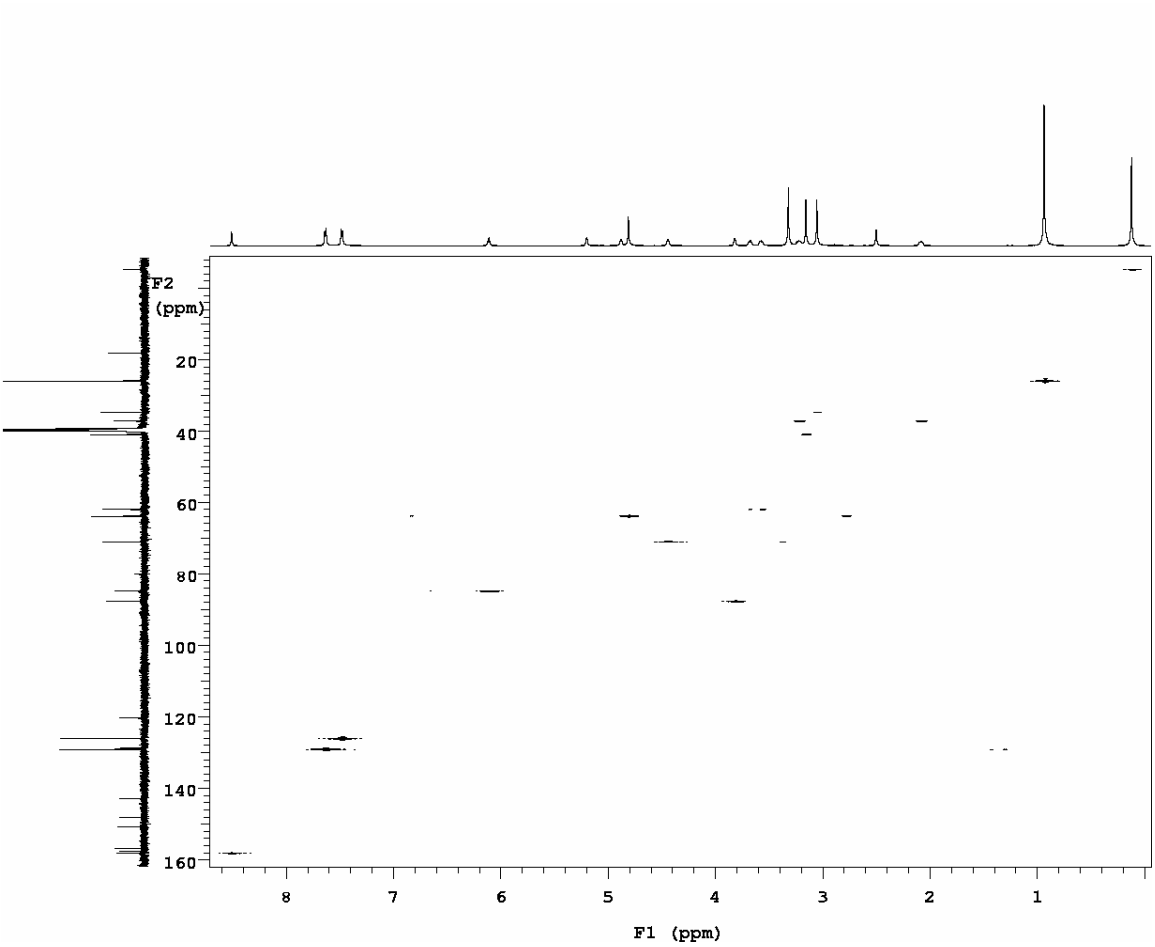
^{13}C NMR of N2-(N,N-Dimethylformamidine)-8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine (9)



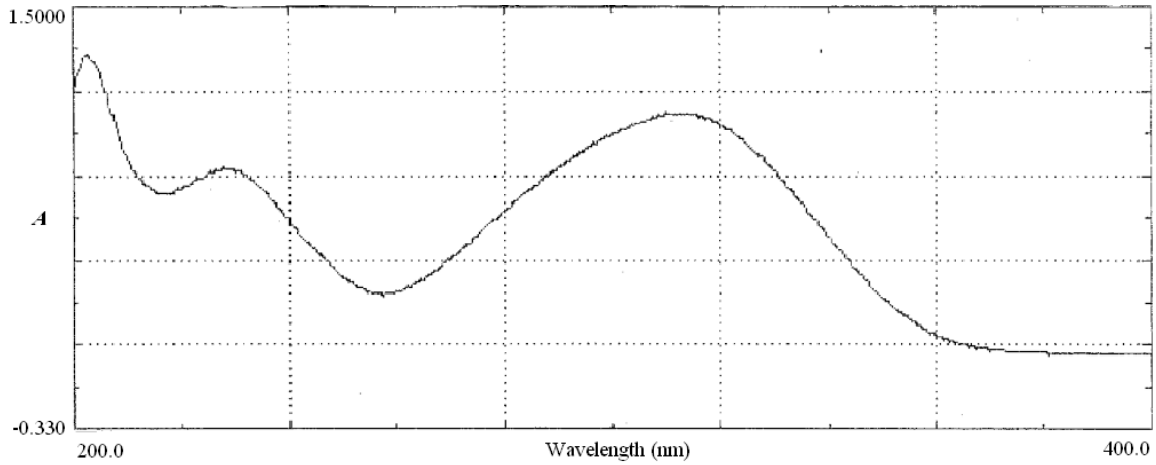
^{1}H-^{1}H COSY NMR of N2-(N,N-Dimethylformamidine)-8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine (9)



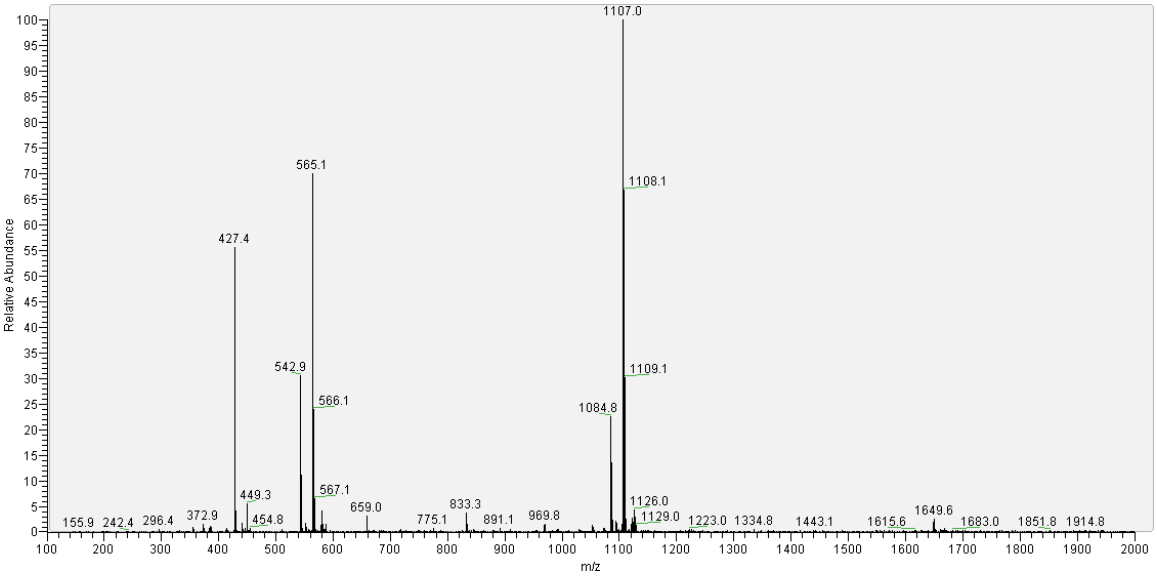
^{1}H-^{13}C HETCOR NMR of N2-(N,N-Dimethylformamidine)-8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine (9)



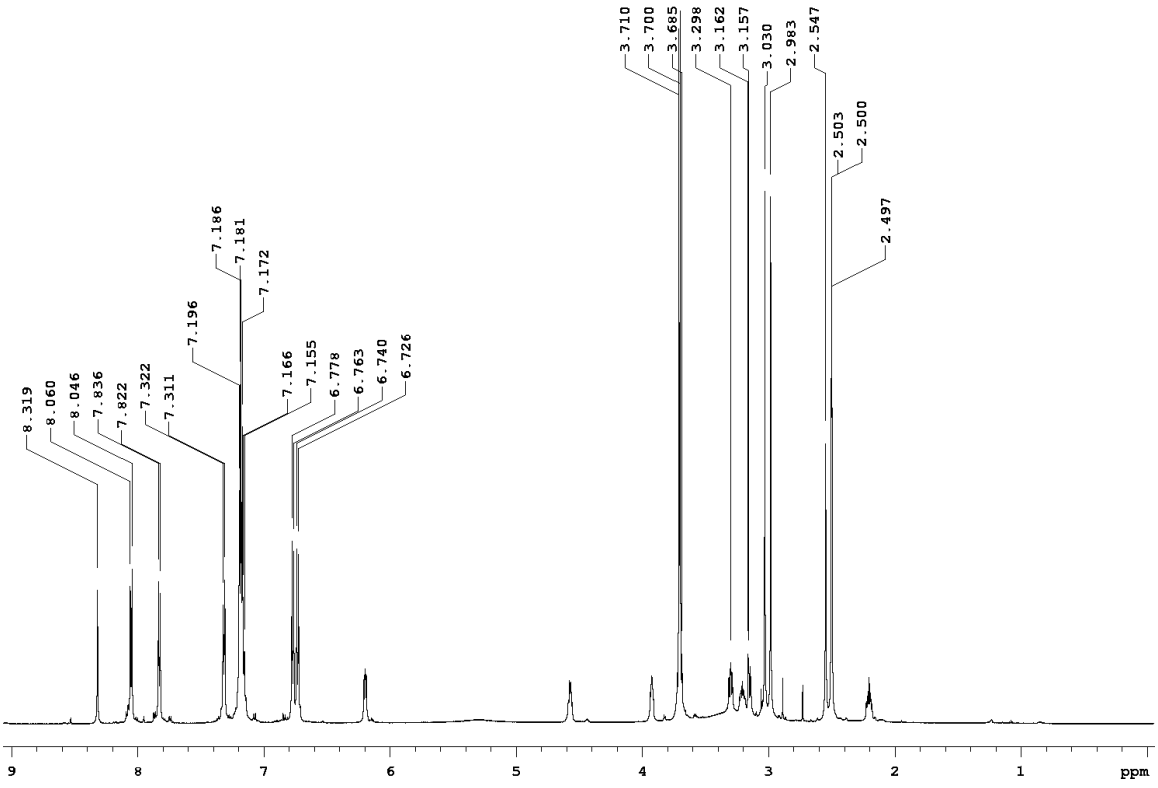
UV Spectrum of N2-(N,N-Dimethylformamidine)-8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine (9)



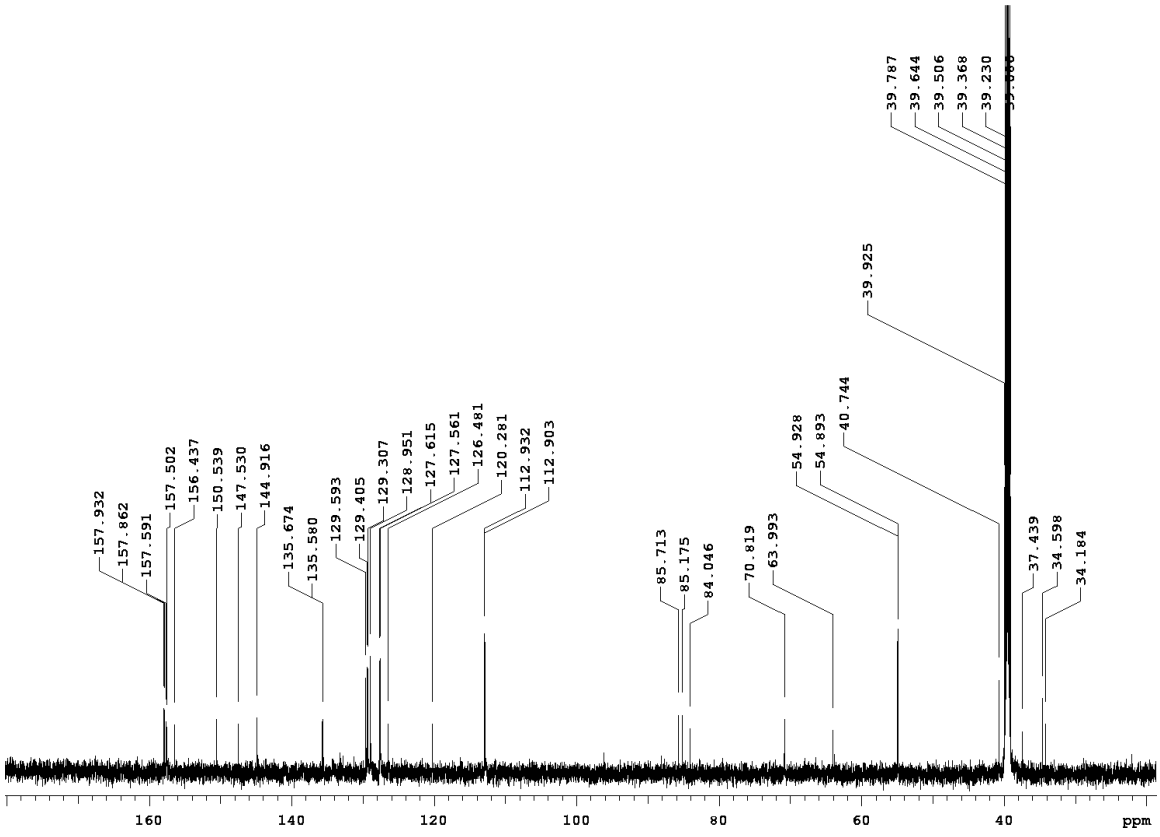
MS$^+$ of N2-(N,N-Dimethylformamidine)-8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine (9)



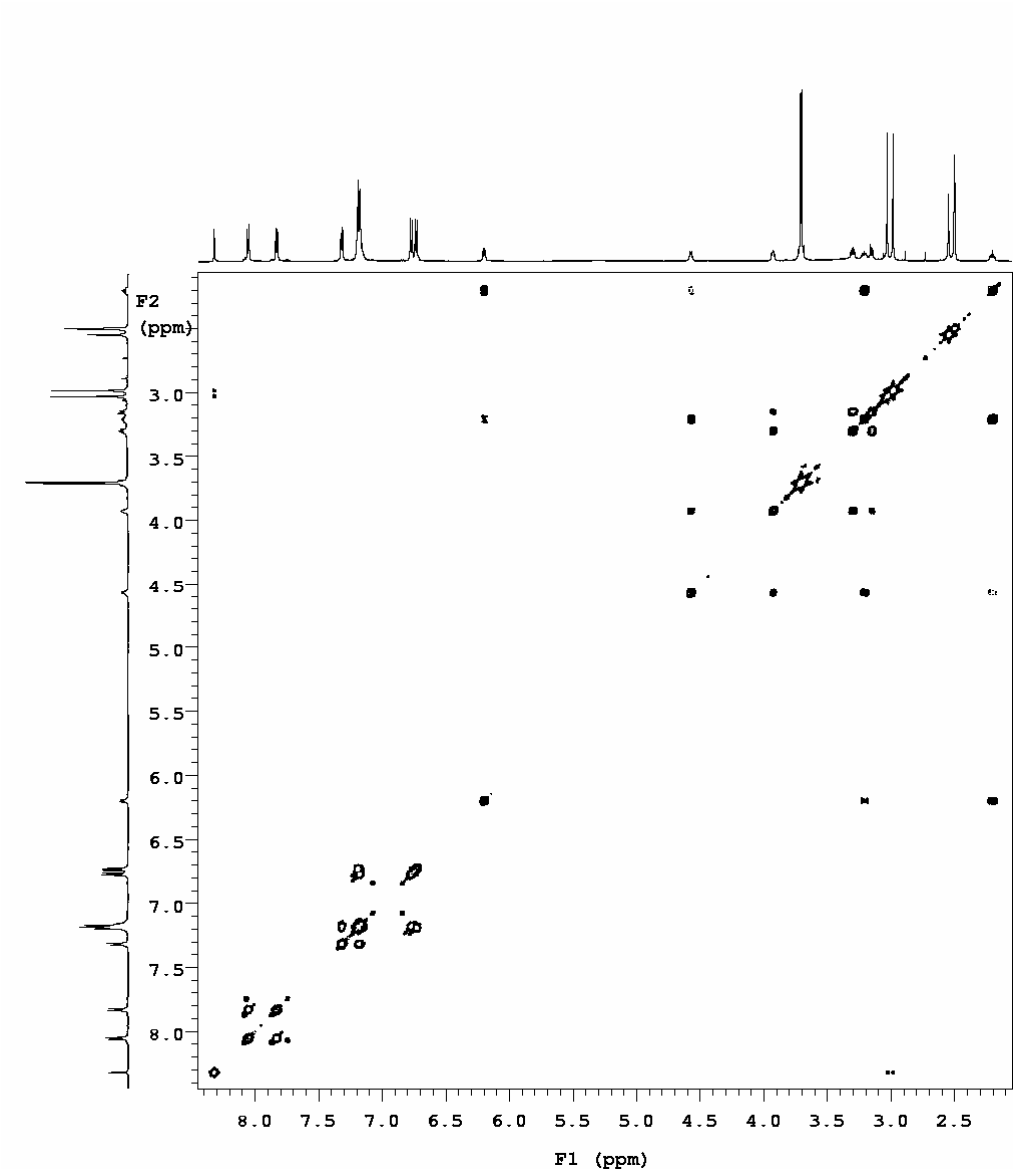
^{1}H NMR of 5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-carboxyphenyl)-2′-deoxyguanosine (10)



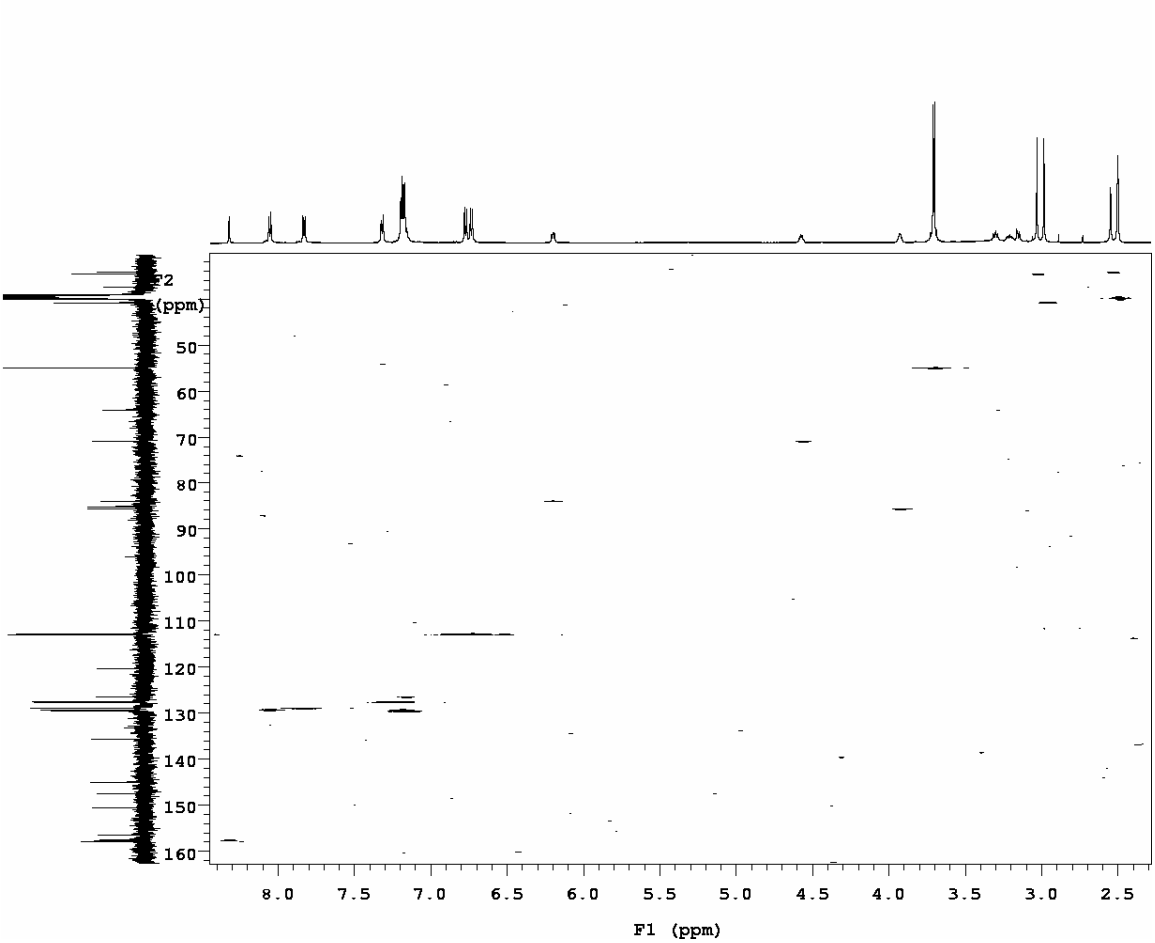
^{13}C NMR of 5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-carboxyphenyl)-2′-deoxyguanosine (10)



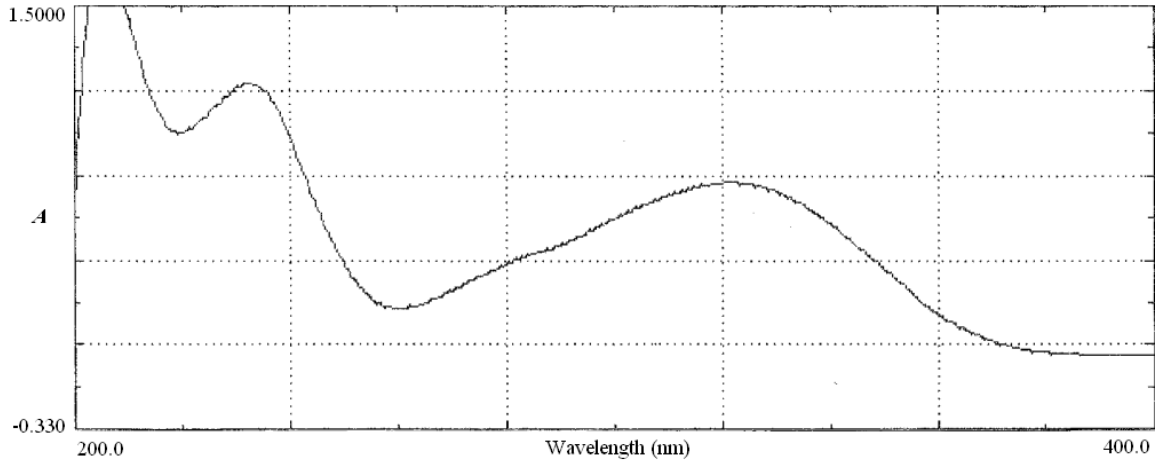
^{1}H-^{1}H COSY NMR of 5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-carboxyphenyl)-2′-deoxyguanosine (10)



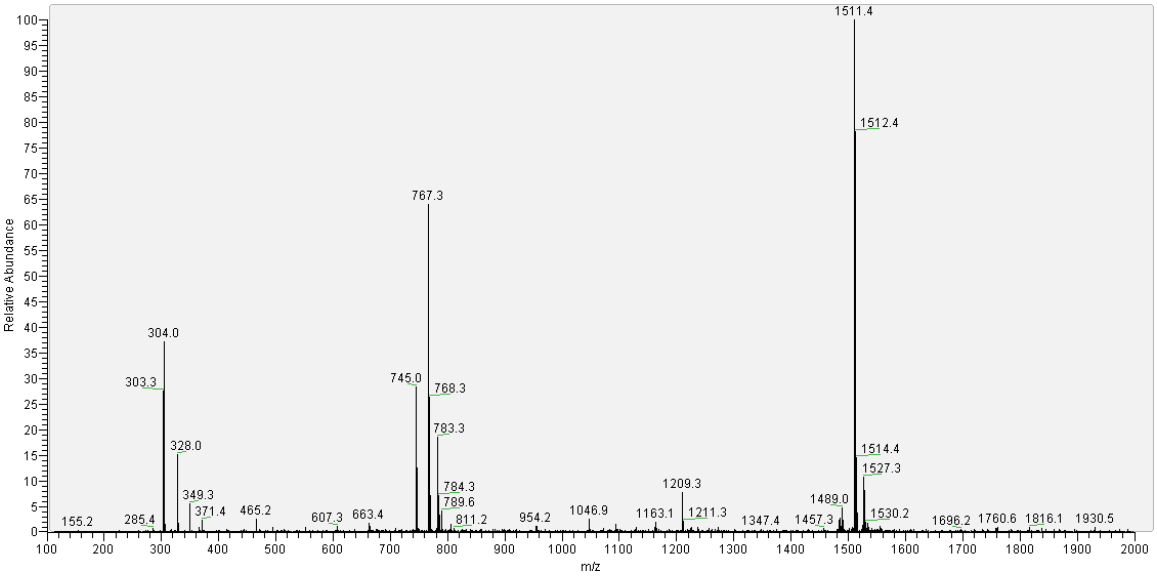
^{1}H-^{13}C HETCOR NMR of 5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-carboxyphenyl)-2′-deoxyguanosine (10)



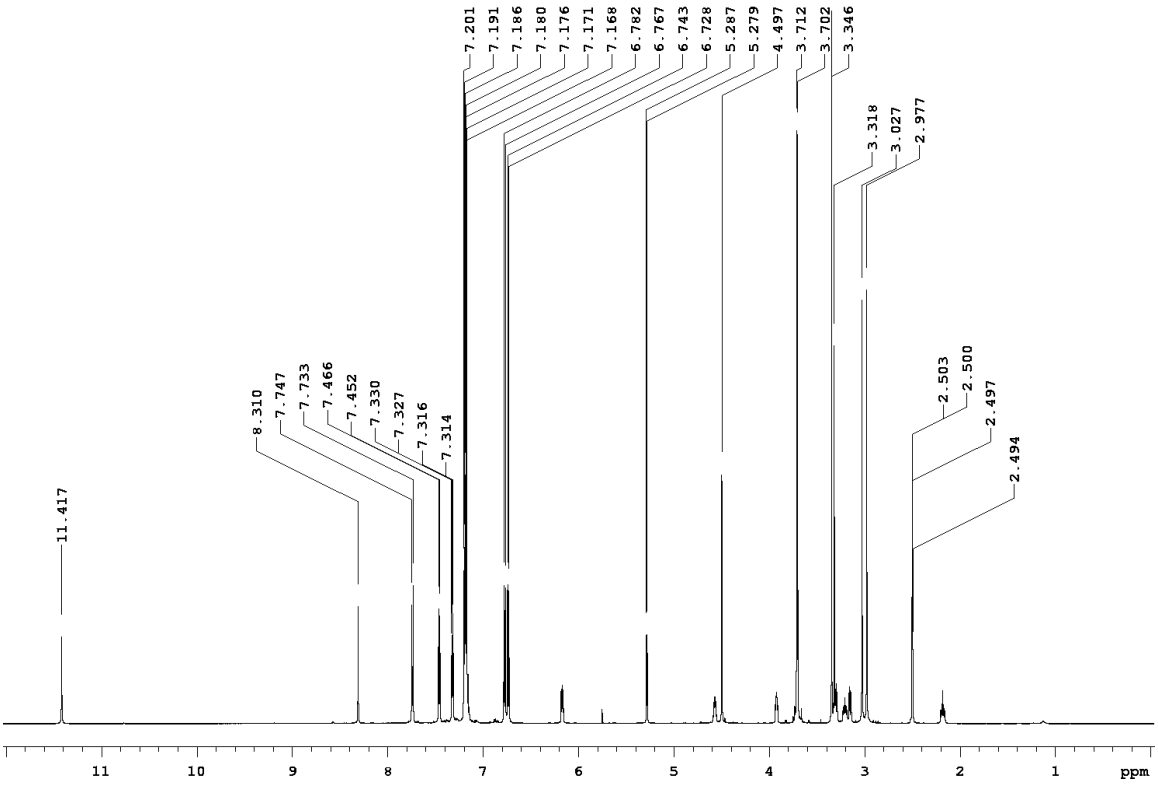
UV Spectrum of 5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-carboxyphenyl)-2′-deoxyguanosine (10)



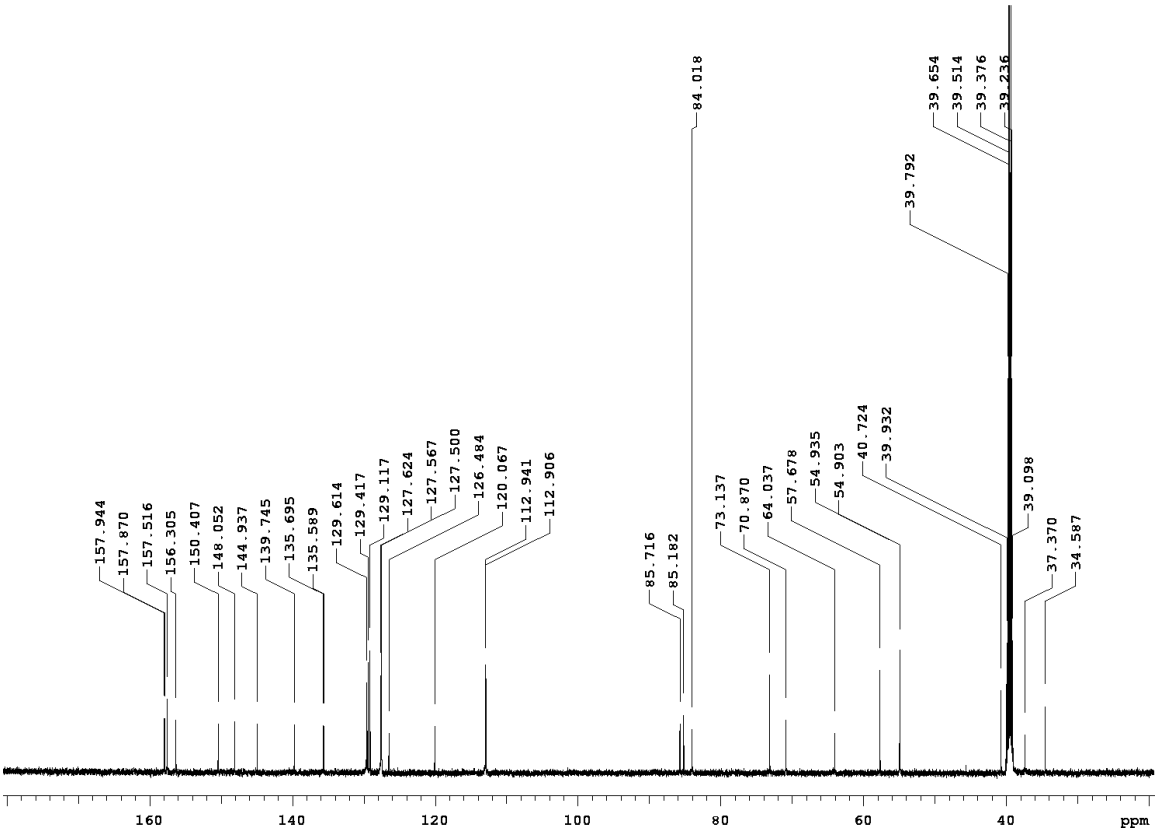
MS$^+$ of 5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-carboxyphenyl)-2′-deoxyguanosine (10)



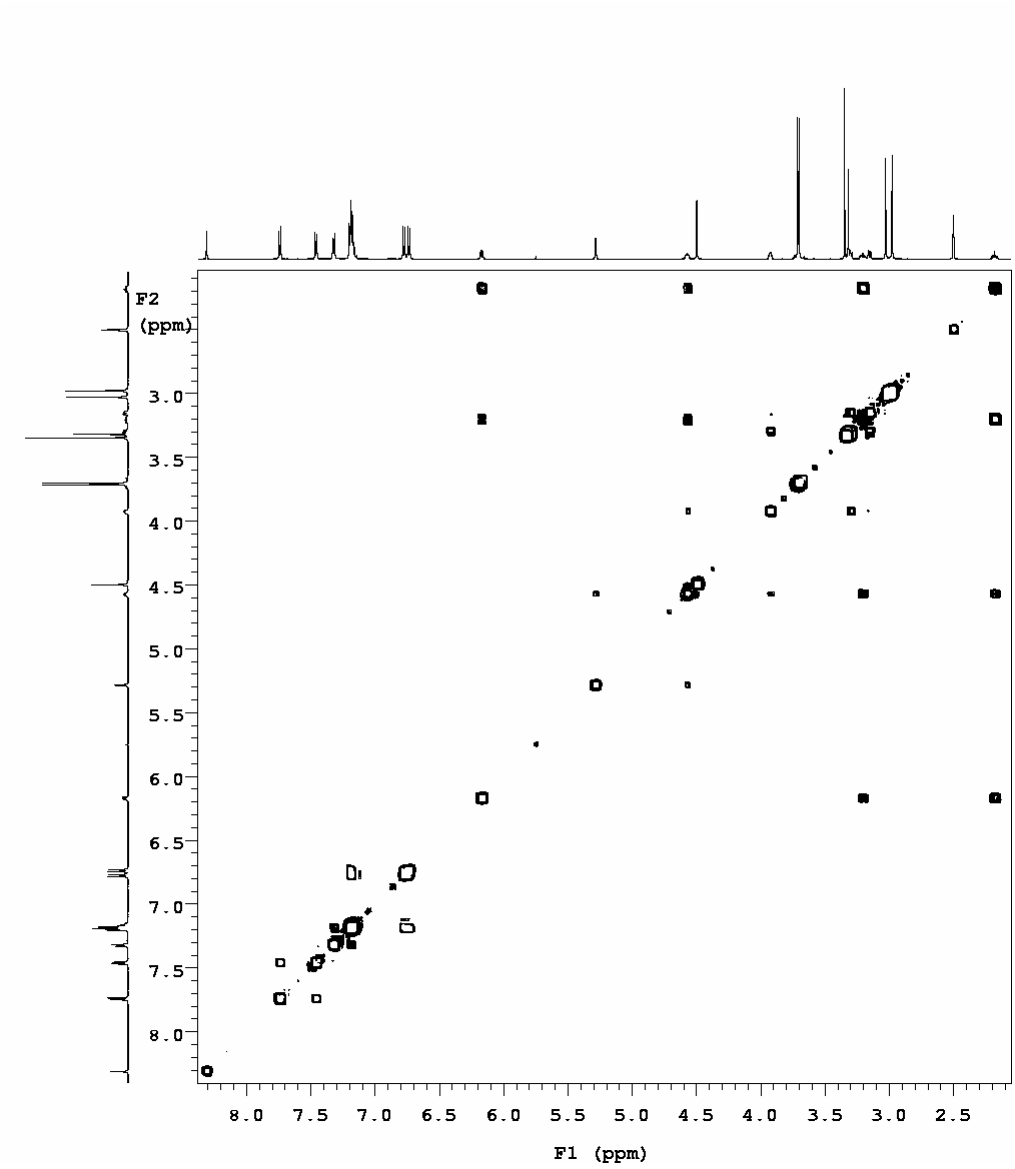
^{1}H NMR of 5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-methoxymethylphenyl)-2′-deoxyguanosine (11)



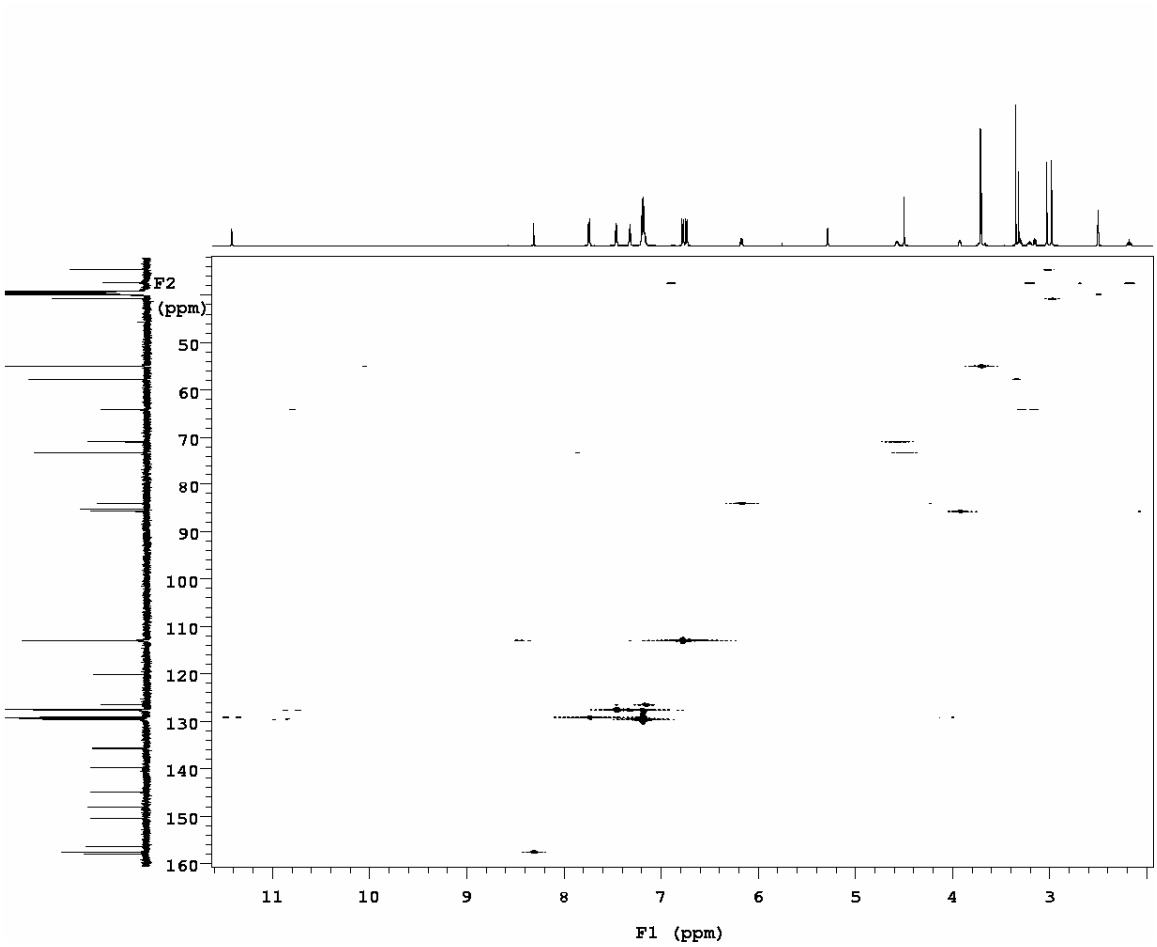
^{13}C NMR of 5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-methoxymethylphenyl)-2′-deoxyguanosine (11)



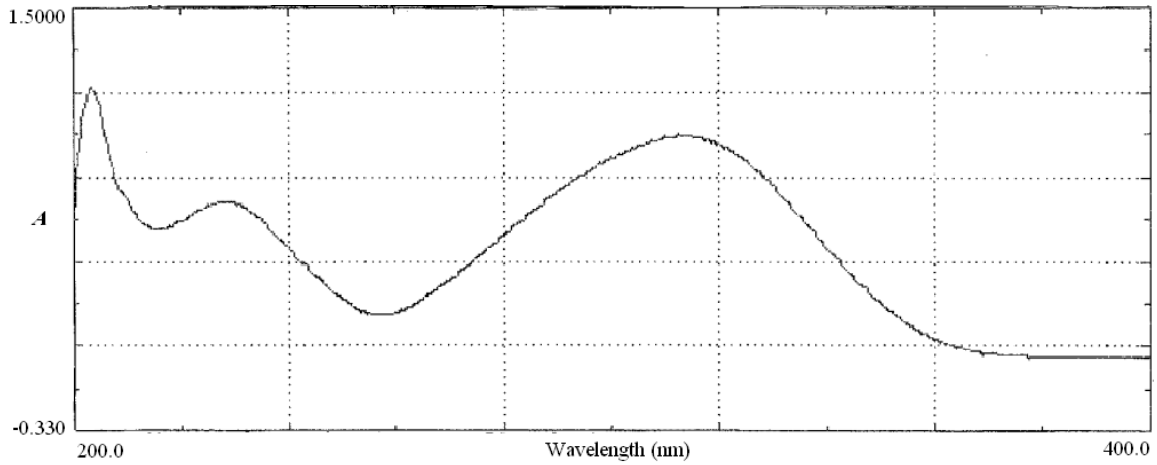
^{1}H-^{1}H COSY NMR of 5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-methoxymethylphenyl)-2′-deoxyguanosine (11)



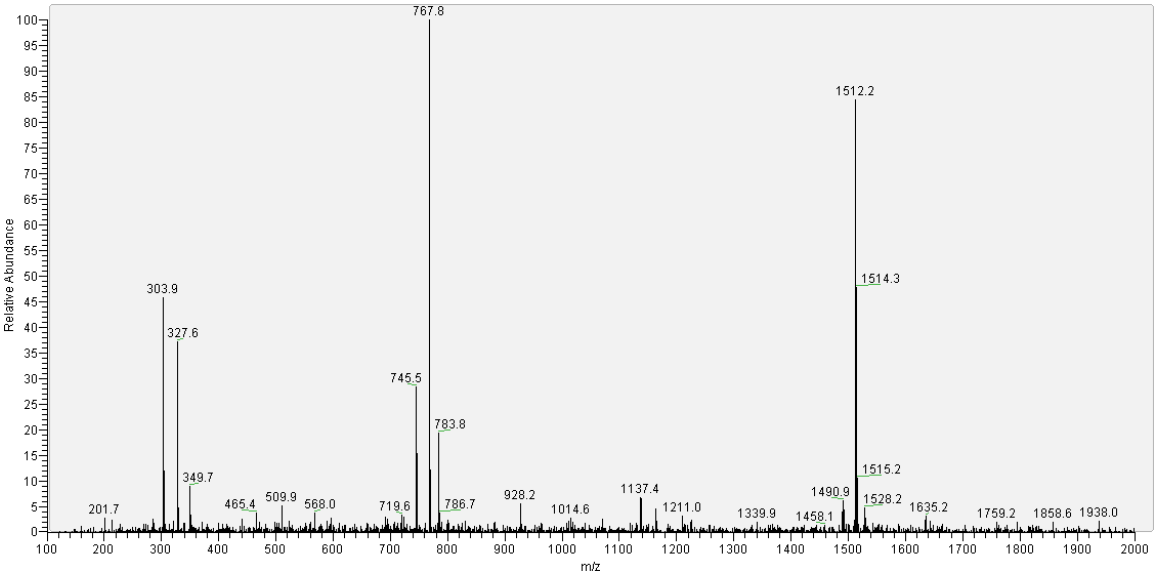
^{1}H-^{13}C HETCOR NMR of 5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-methoxymethylphenyl)-2′-deoxyguanosine (11)



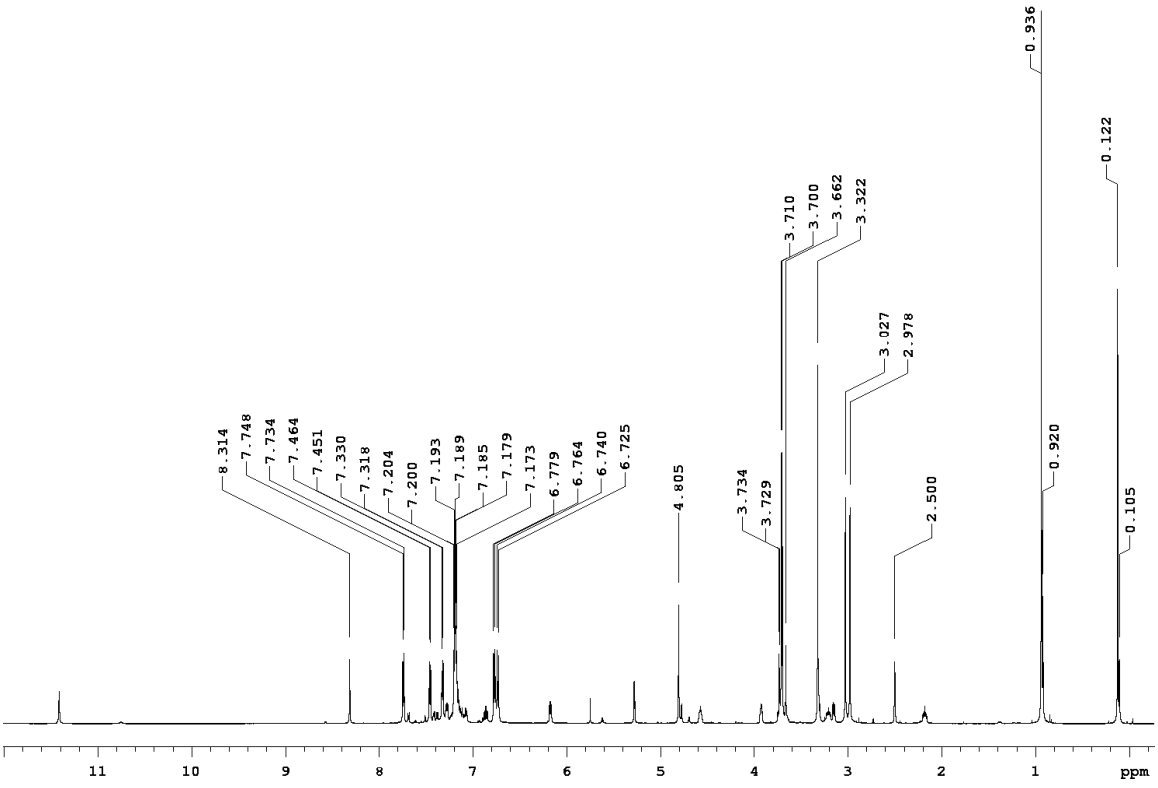
UV Spectrum of 5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-methoxymethylphenyl)-2′-deoxyguanosine (11)



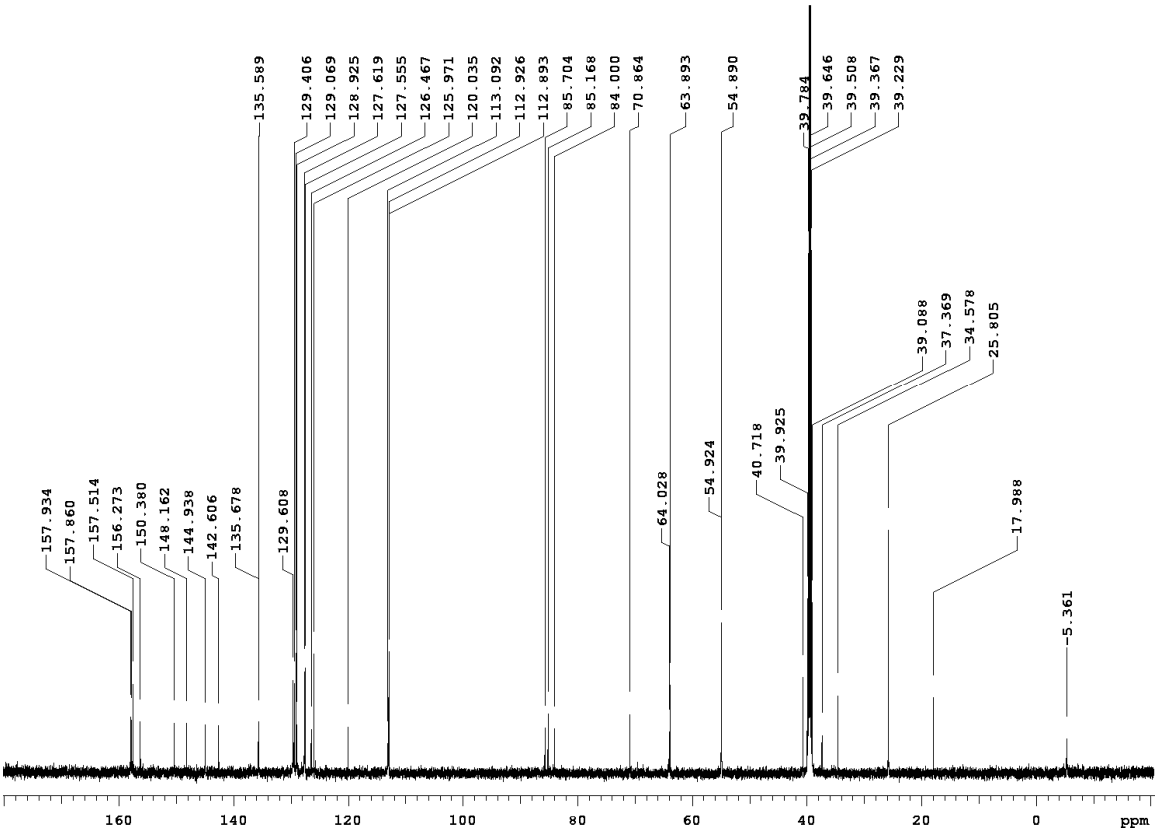
MS$^+$ of 5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-methoxymethylphenyl)-2′-deoxyguanosine (11)



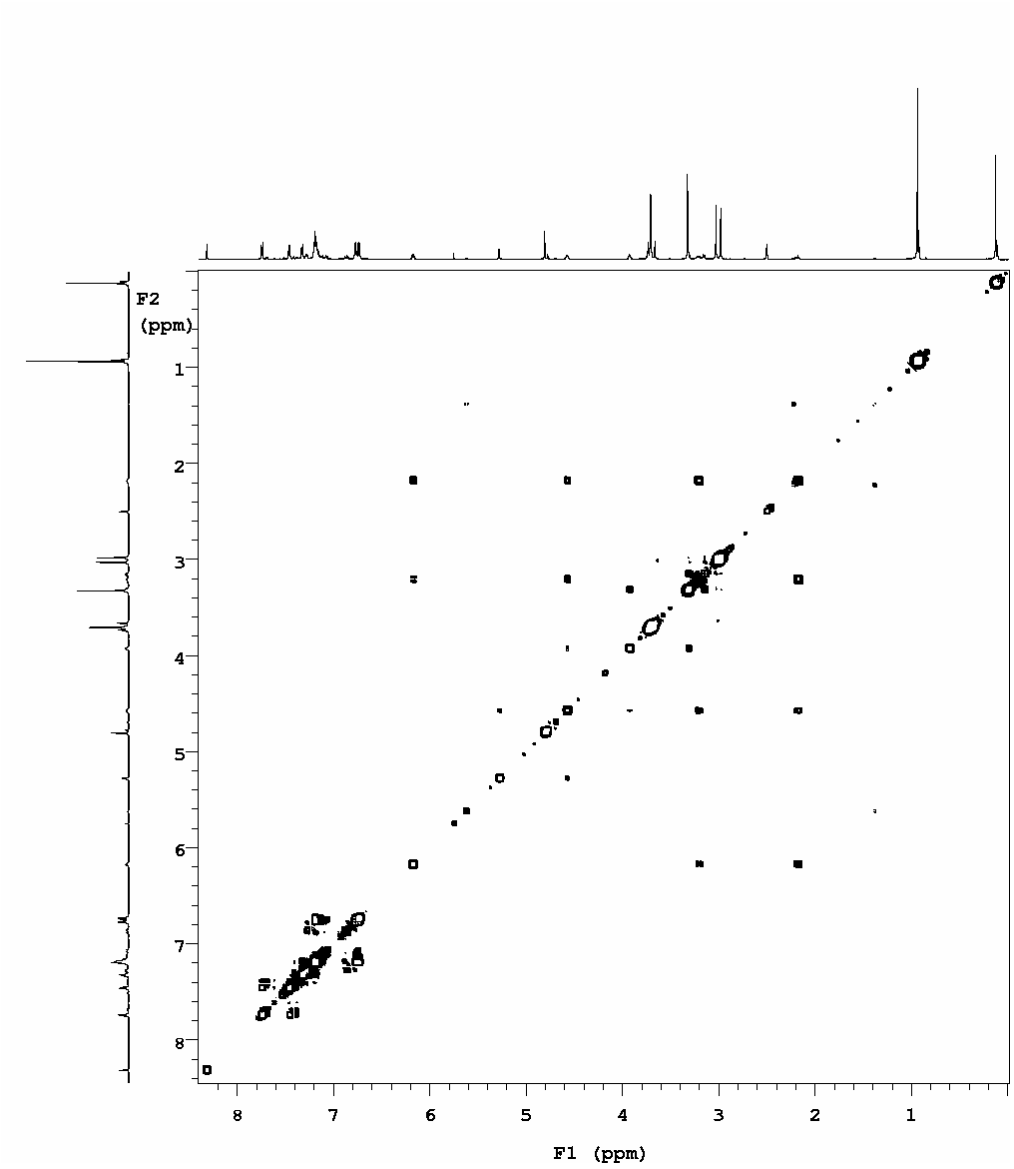
^{1}H NMR of 5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine (12)



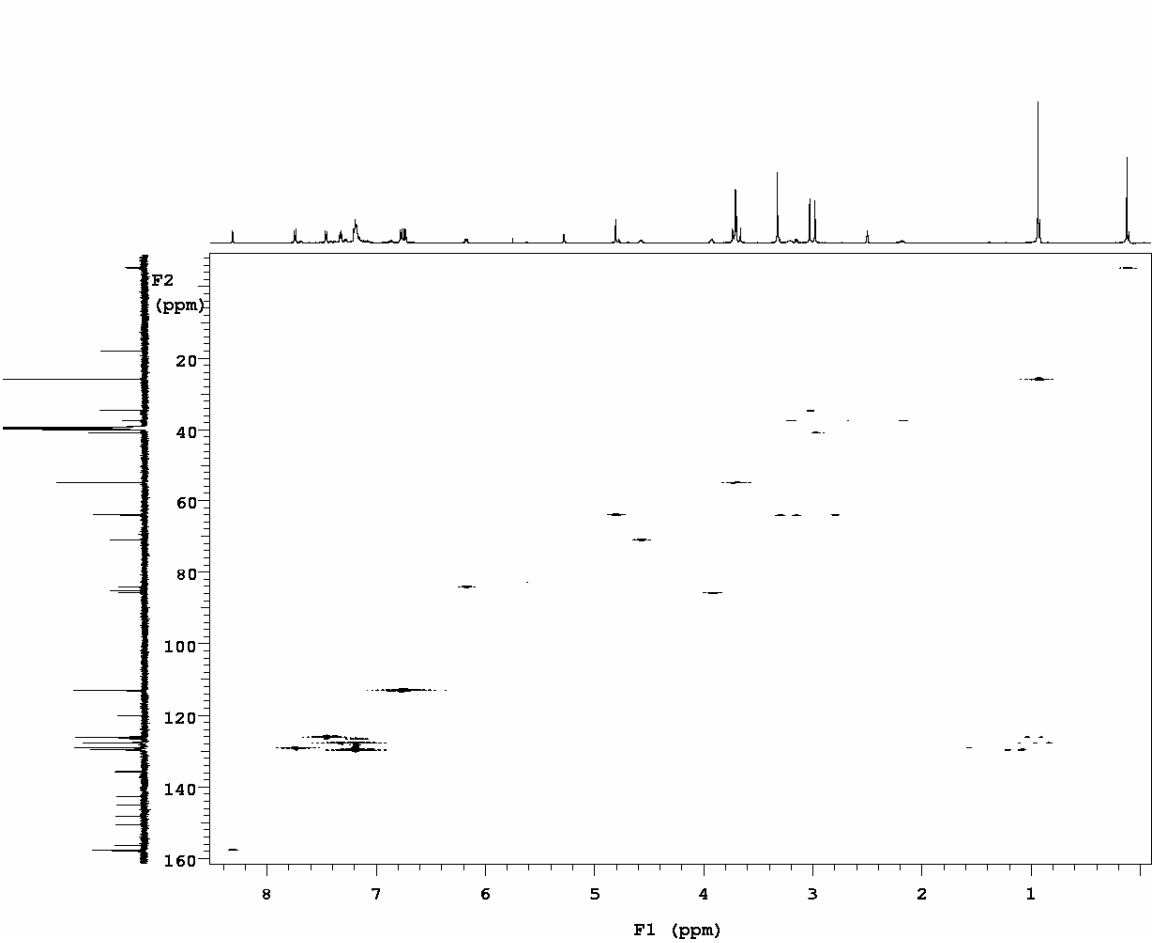
^{13}C NMR of 5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine (12)



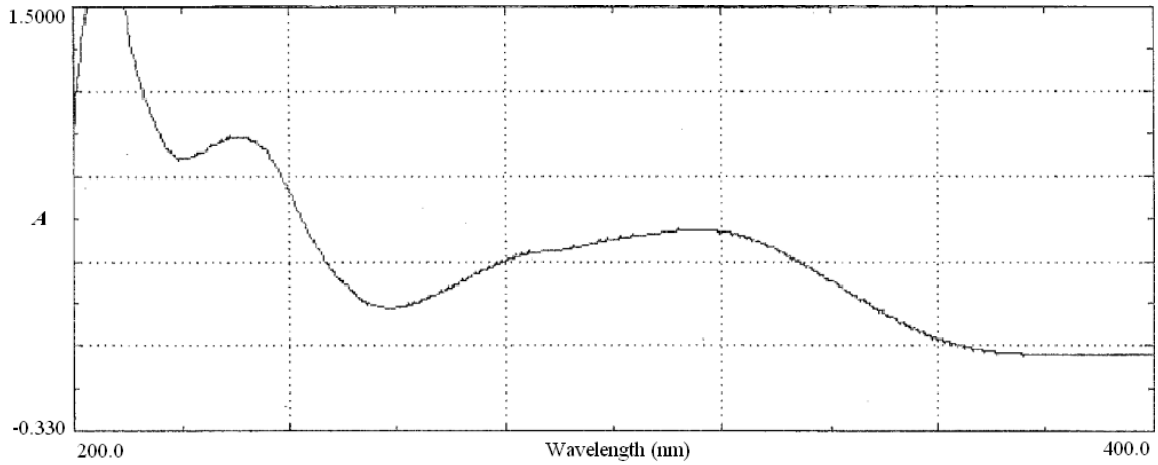
^{1}H-^{1}H COSY NMR of 5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine (12)



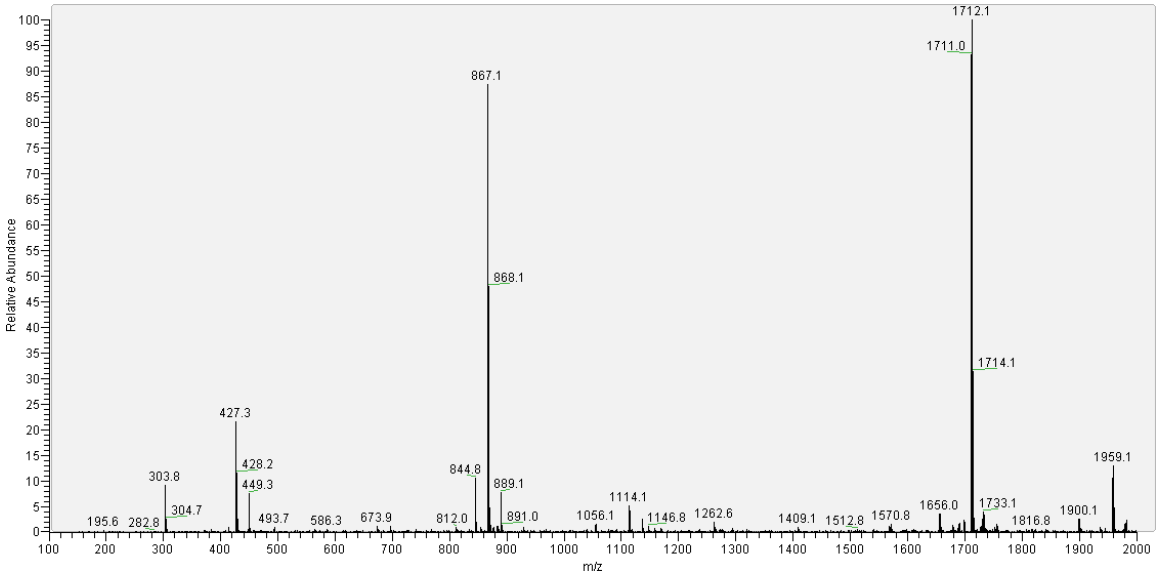
^{1}H-^{13}C HETCOR NMR of 5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine (12)



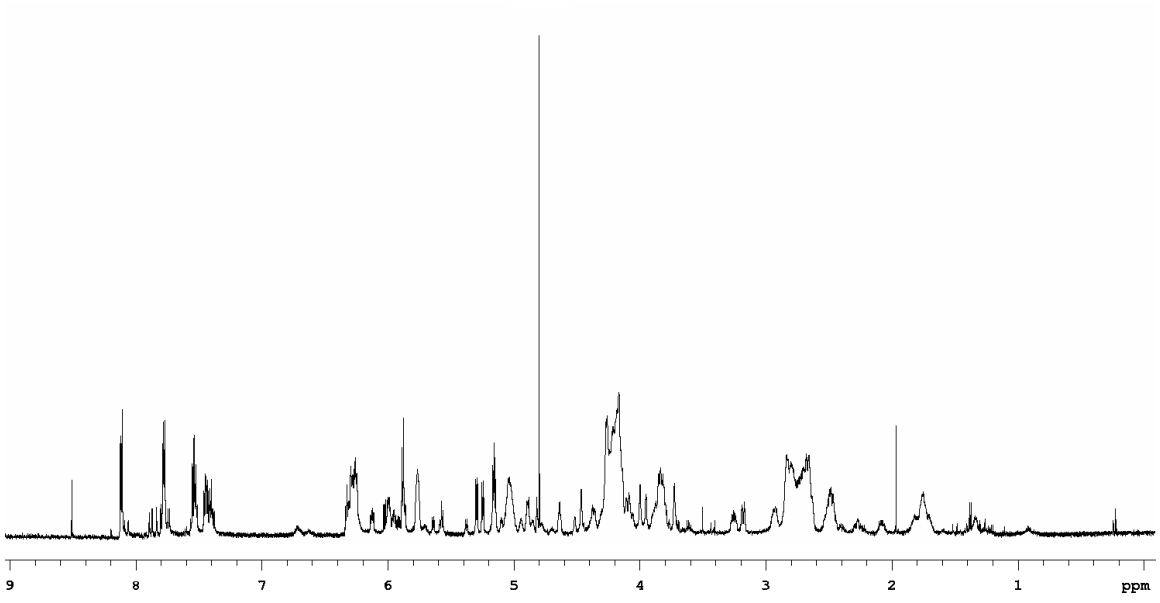
UV Spectrum of 5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine (12)



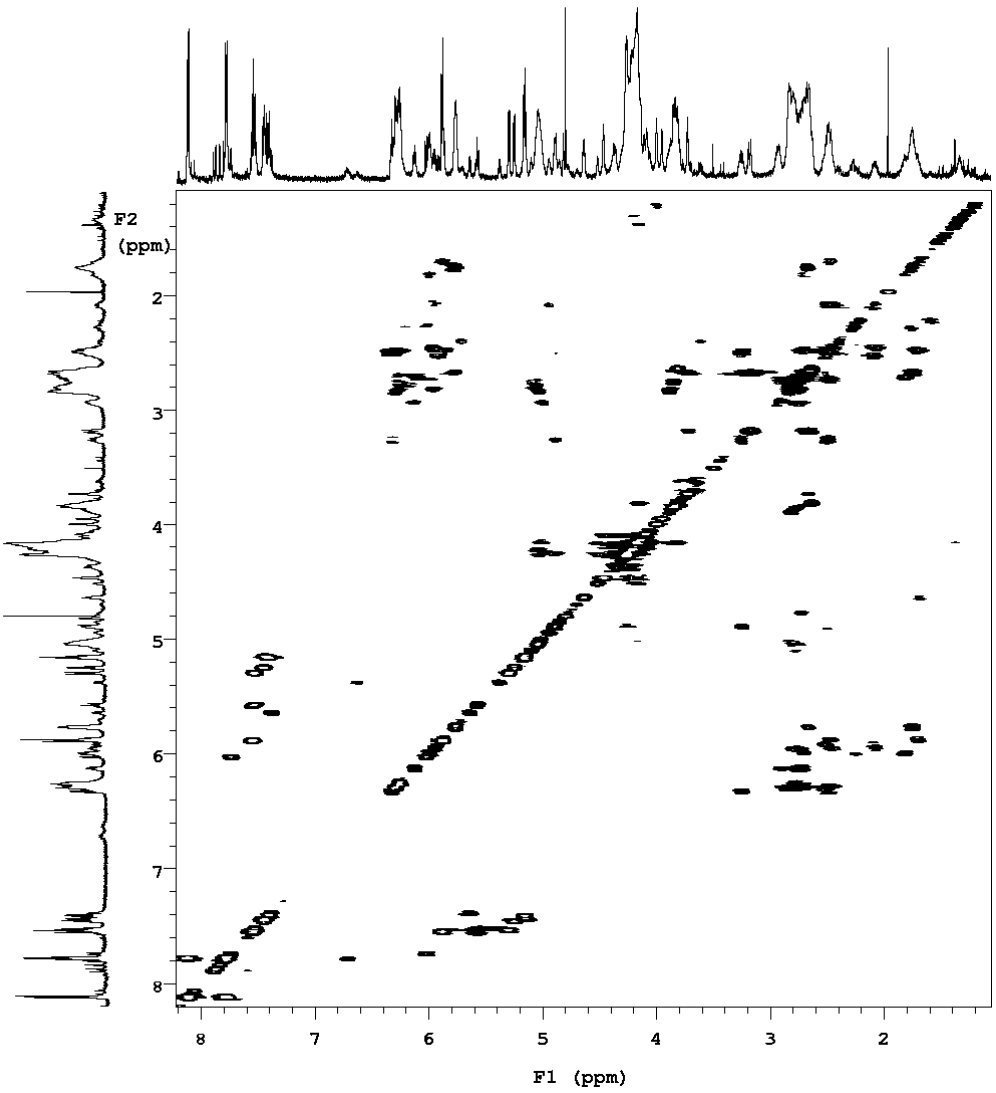
MS$^+$ of 5′-O-(DMTr)-N2-(N,N-dimethylformamidine)-8-(4-(TBS-O-methyl)phenyl)-2′-deoxyguanosine (12)



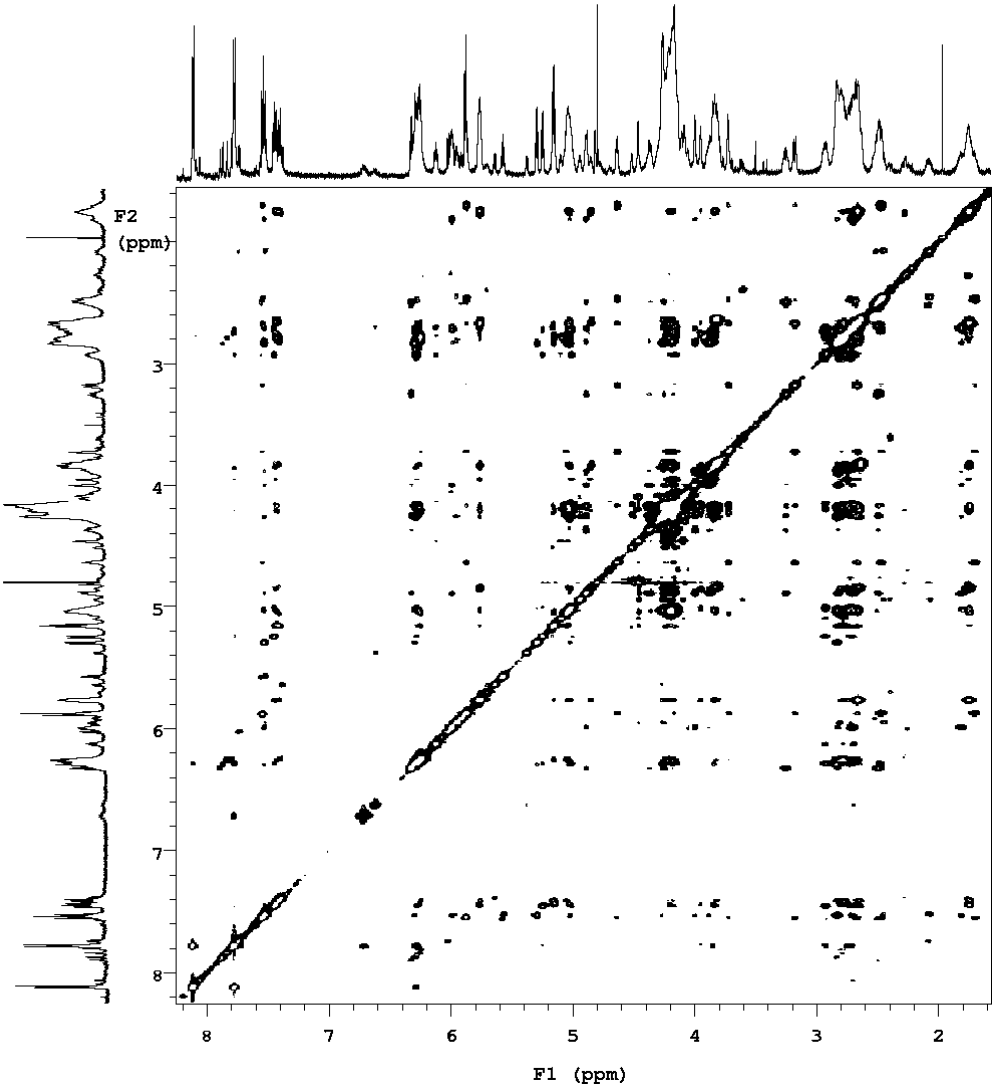
^{1}H NMR of CG^{8CPh} in 10 mM Phosphate Buffer pD 7.4 at 28°C



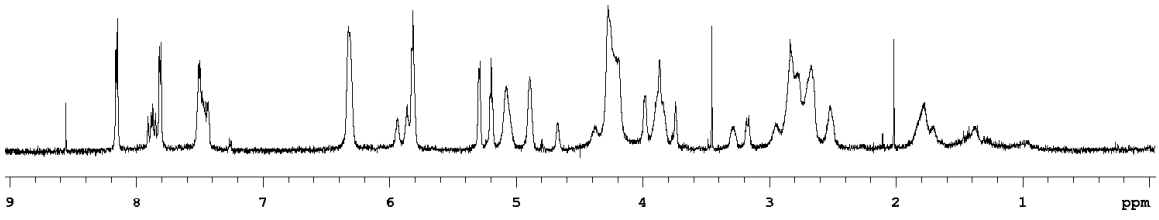
^{1}H-^{1}H COSY NMR of CG^{8CPh} in 10 mM Phosphate Buffer pD 7.4 at 28°C



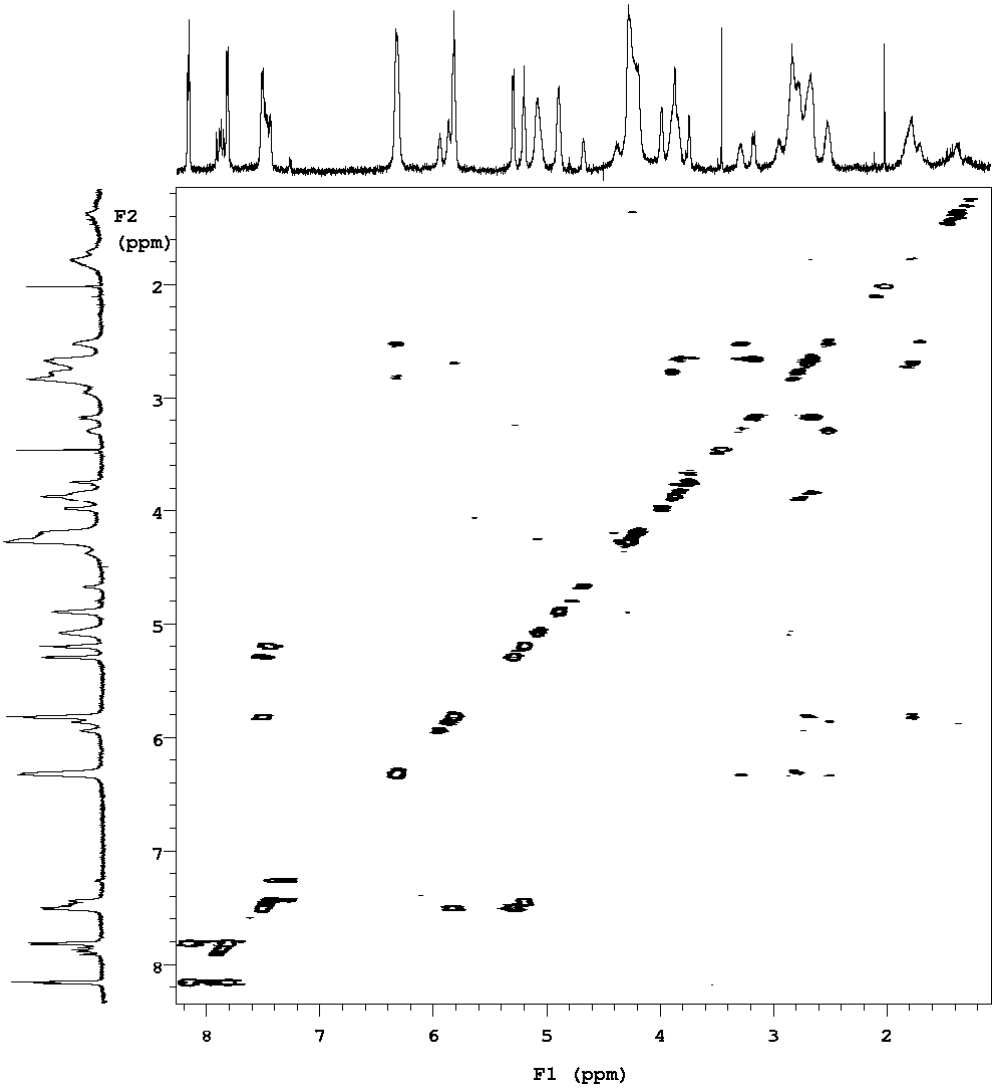
^{1}H-^{1}H NOESY NMR of CG^{8CPh} in 10 mM Phosphate Buffer pD 7.4 at 28°C



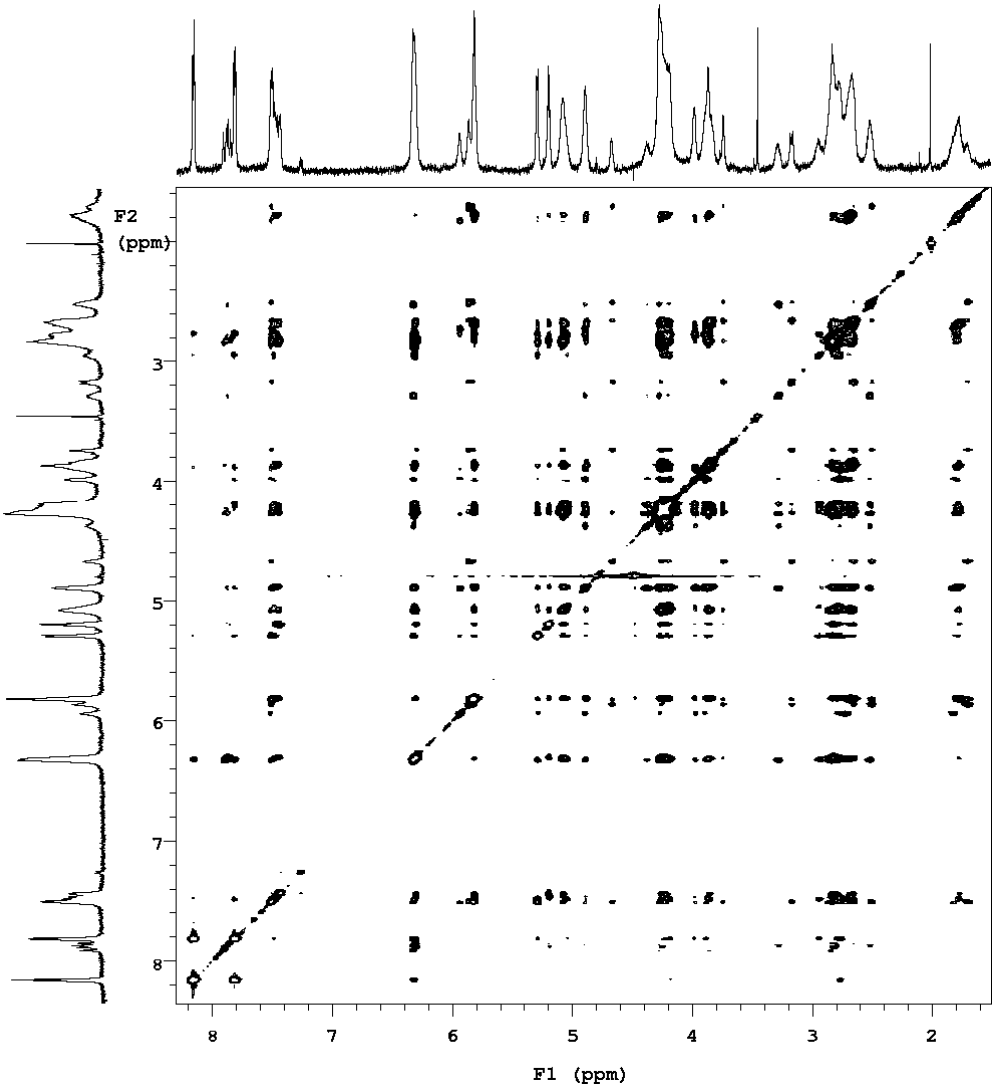
^{1}H NMR of CG^{8CPh} in 10 mM Phosphate Buffer pD 7.4 and 500 mM NaCl at 28°C



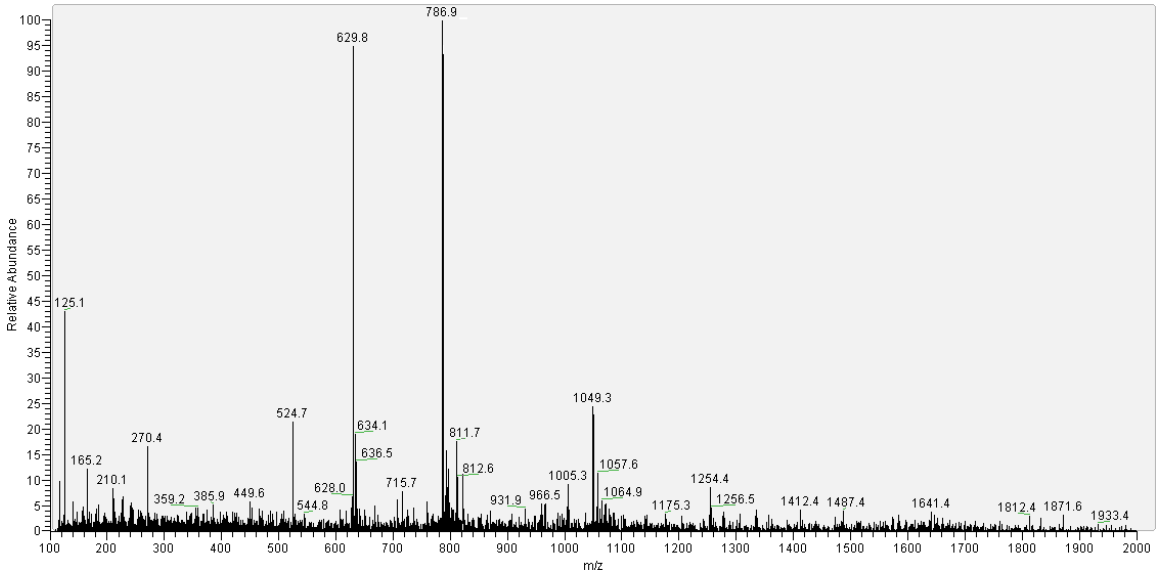
^{1}H-^{1}H COSY NMR of CG^{8CPh} in 10 mM Phosphate Buffer pD 7.4 and 500 mM NaCl at 28°C



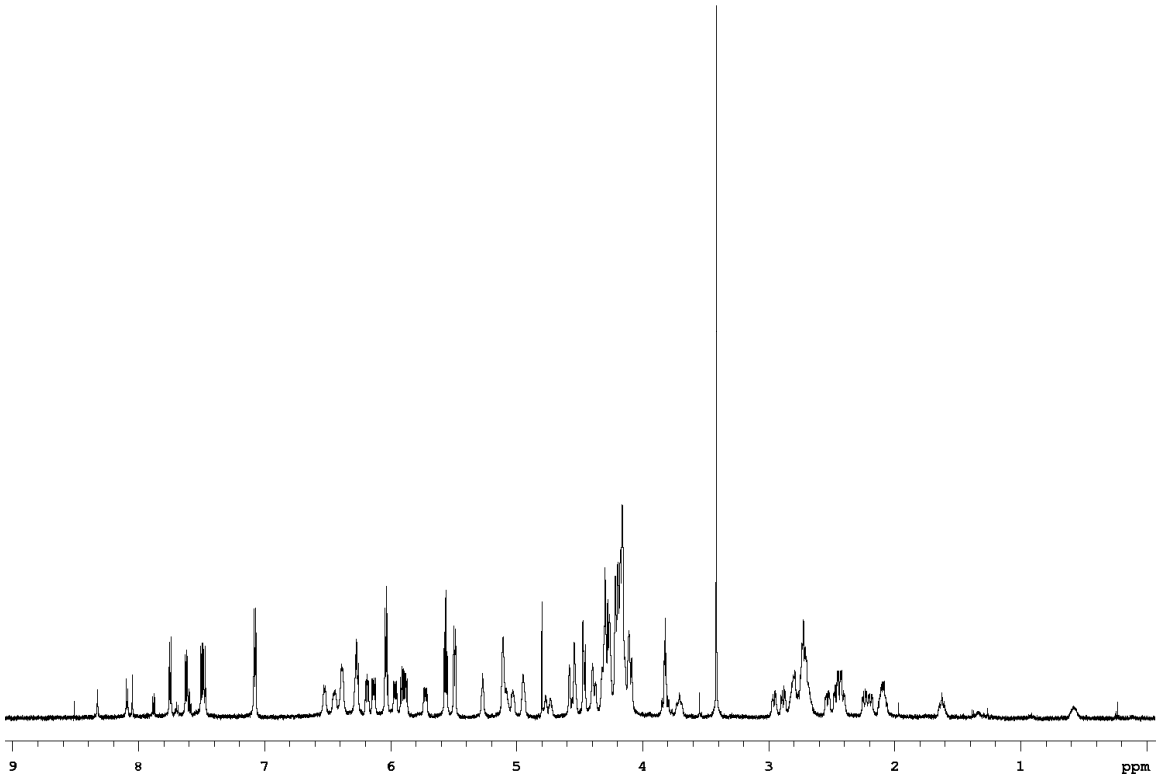
^{1}H-^{1}H NOESY NMR of CG^{8CPh} in 10 mM Phosphate Buffer pD 7.4 and 500 mM NaCl at 28°C



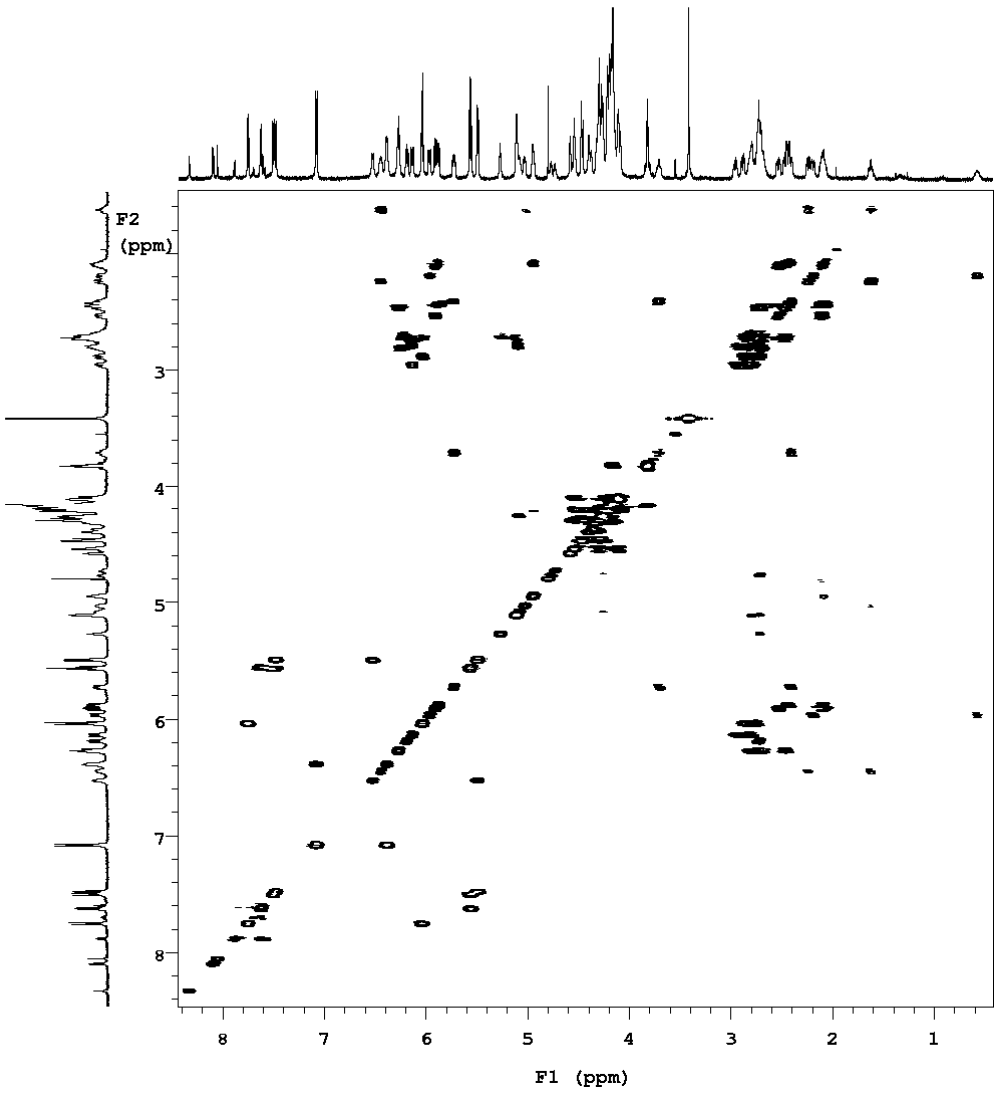
MS^{-} of CG^{8CPh}


Relative Abundance
m/z
786.9
629.8
125.1
1049.3
524.7
634.1
811.7
270.4
636.5
165.2
812.6
1057.6
210.1
1005.3
1254.4
715.7
1064.9
359.2
385.9
449.6
628.0
544.8
931.9
966.5
1175.3
1256.5
1412.4
1487.4
1641.4
1812.4
1871.6
1933.4

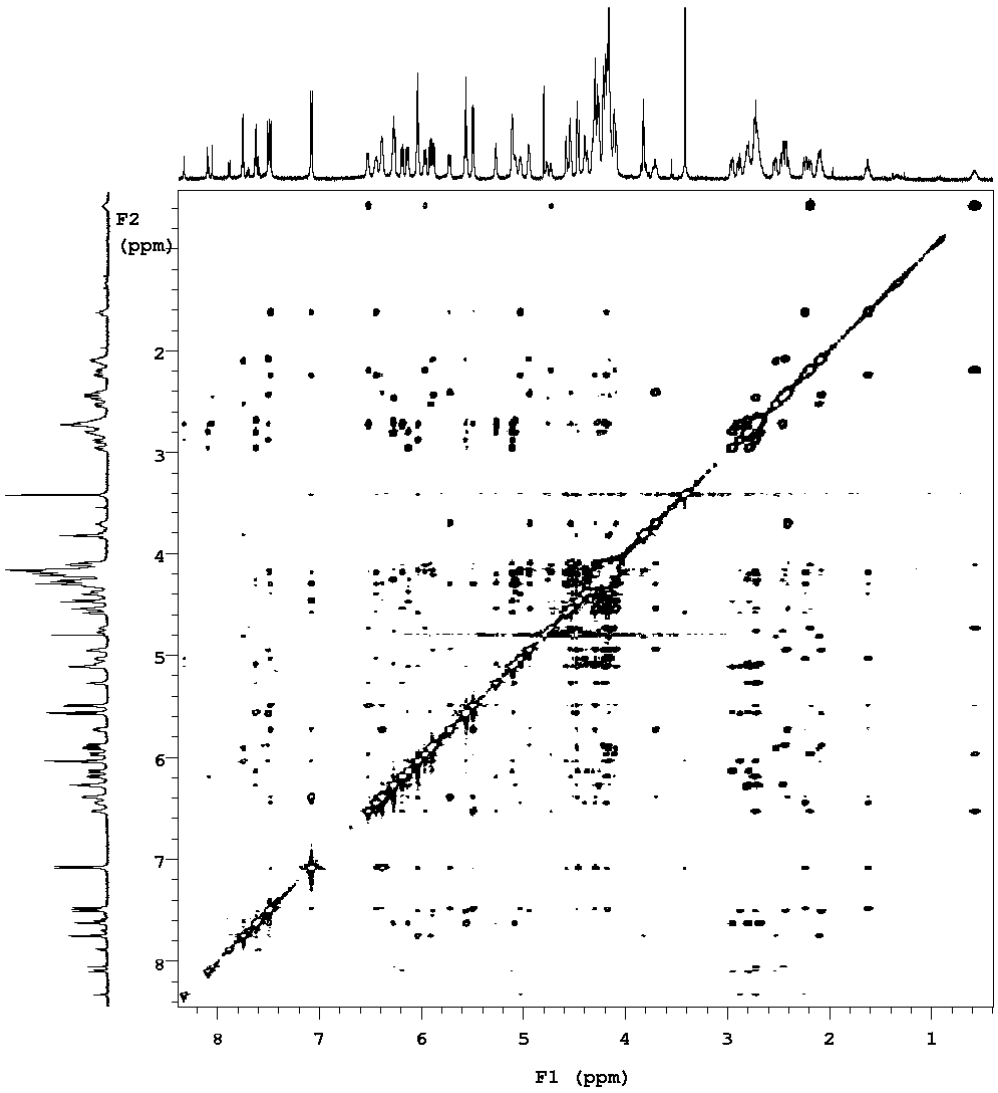
^{1}H NMR of CG^{8MMPh} in 10 mM Phosphate Buffer pD 7.4 at 28°C



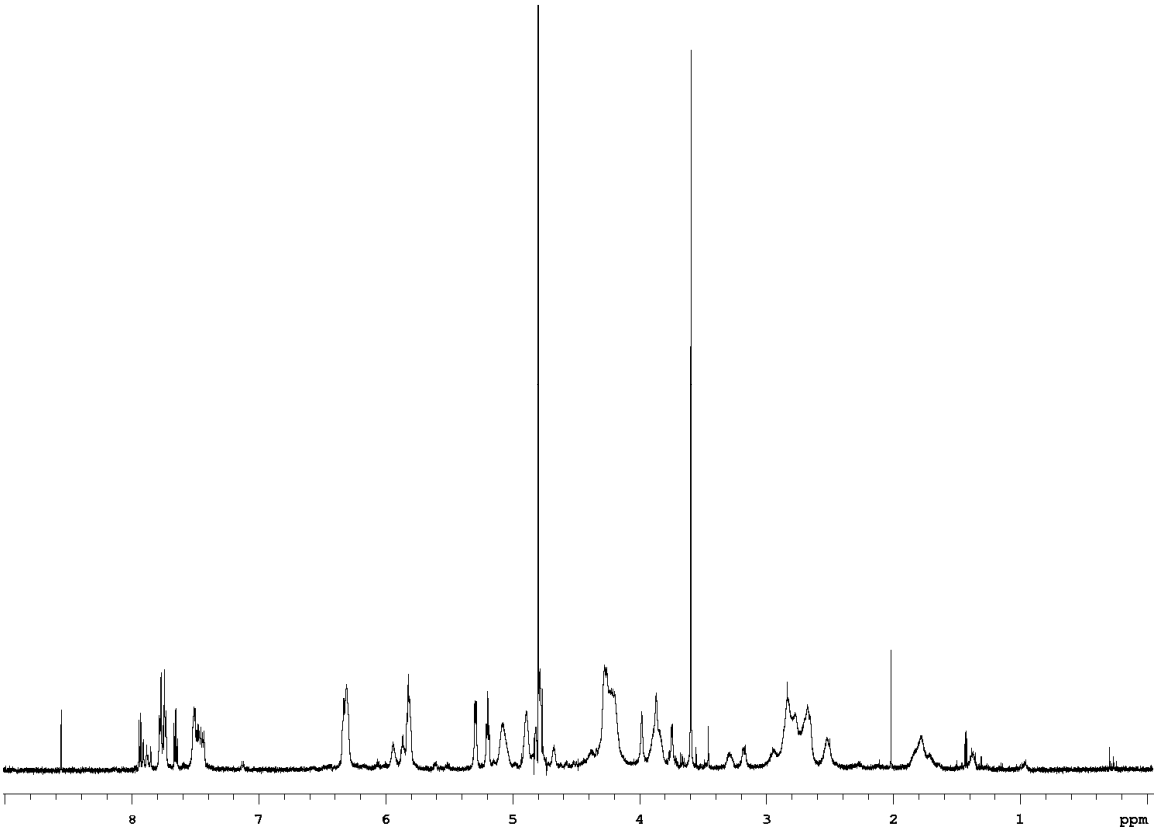
^{1}H-^{1}H COSY NMR of CG^{8MMPh} in 10 mM Phosphate Buffer pD 7.4 at 28°C



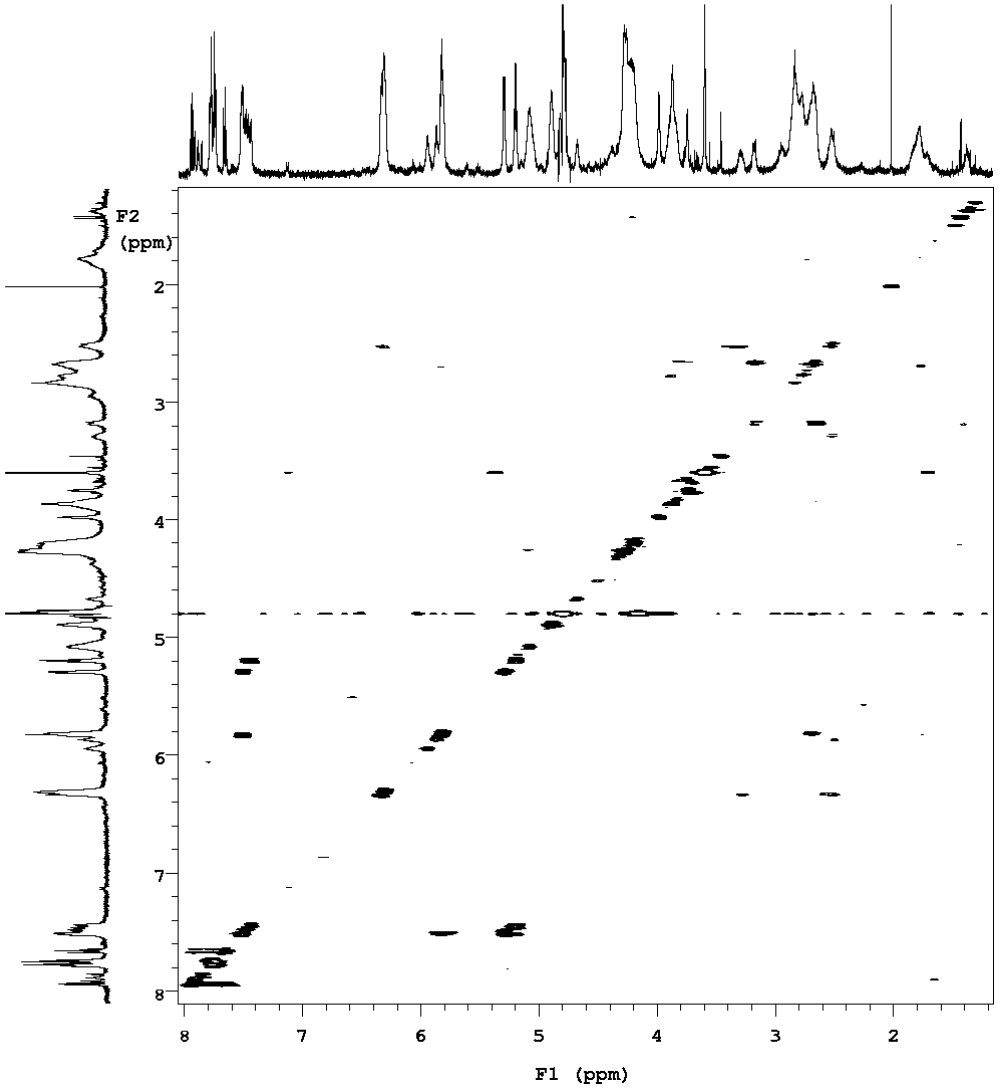
^{1}H-^{1}H NOESY NMR of CG^{8MMPh} in 10 mM Phosphate Buffer pD 7.4 at 28°C



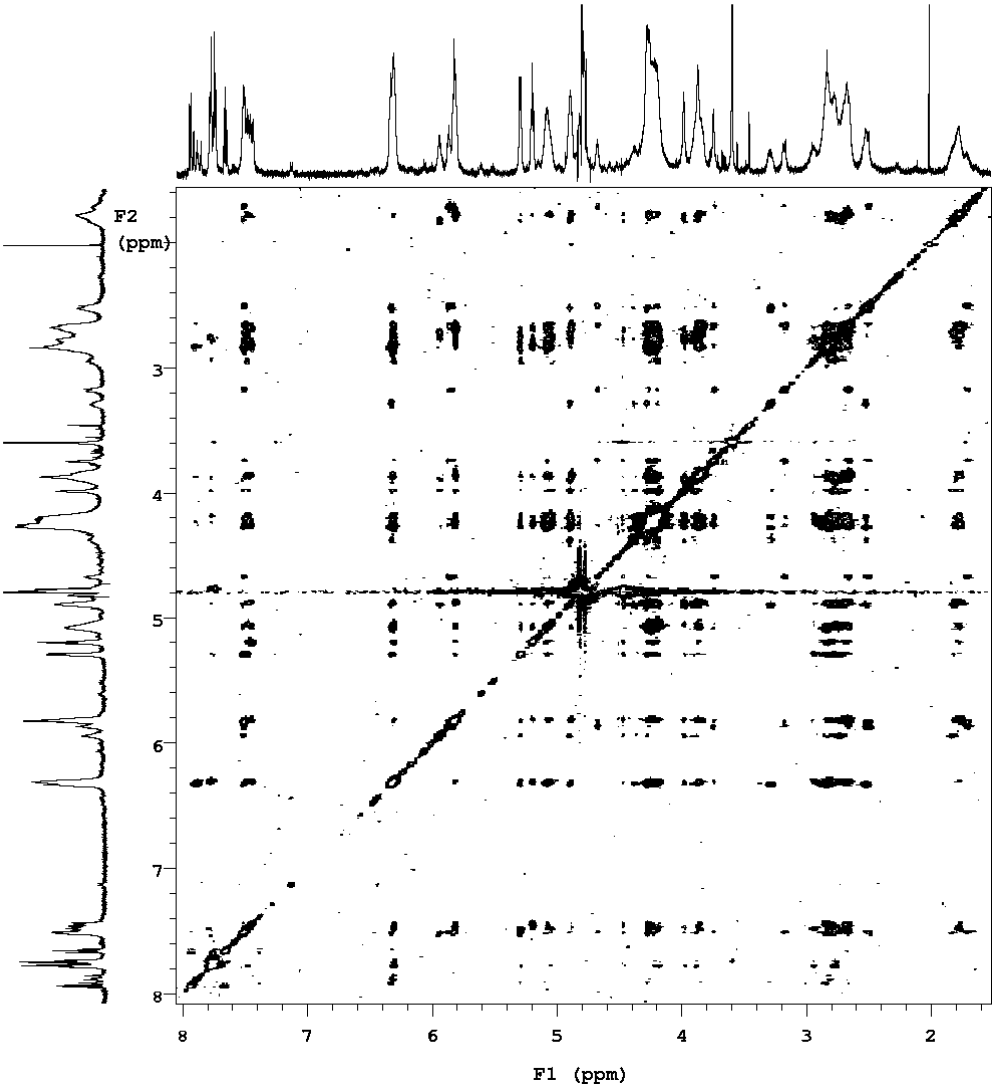
^{1}H NMR of CG^{8MMPh} in 10 mM Phosphate Buffer pD 7.4 and 500 mM NaCl at 28°C



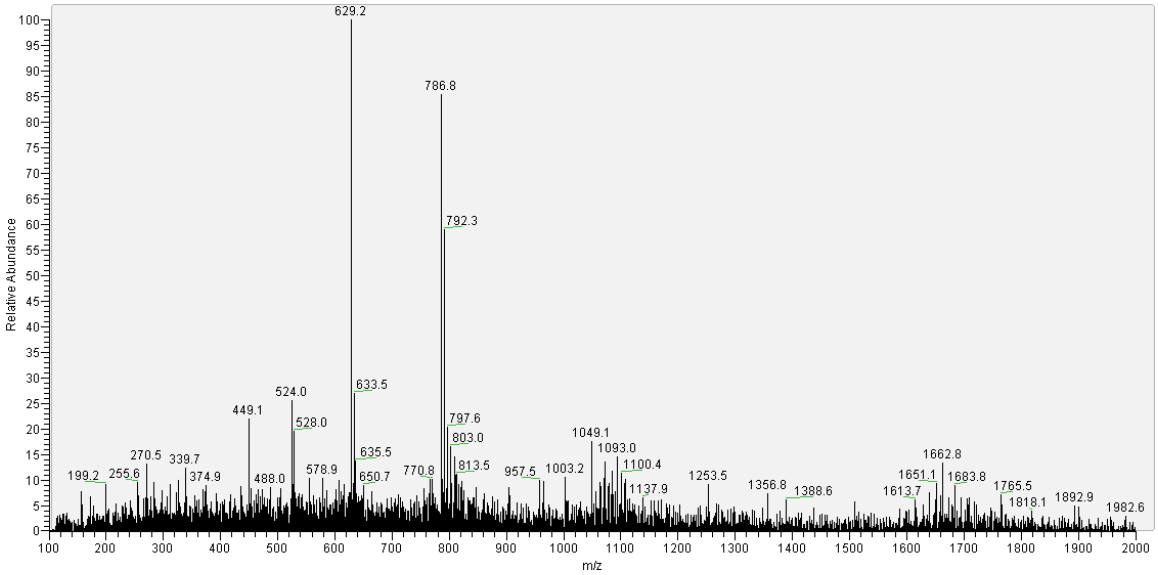
^{1}H-^{1}H COSY NMR of CG^{8MMPh} in 10 mM Phosphate Buffer pD 7.4 and 500 mM NaCl at 28°C



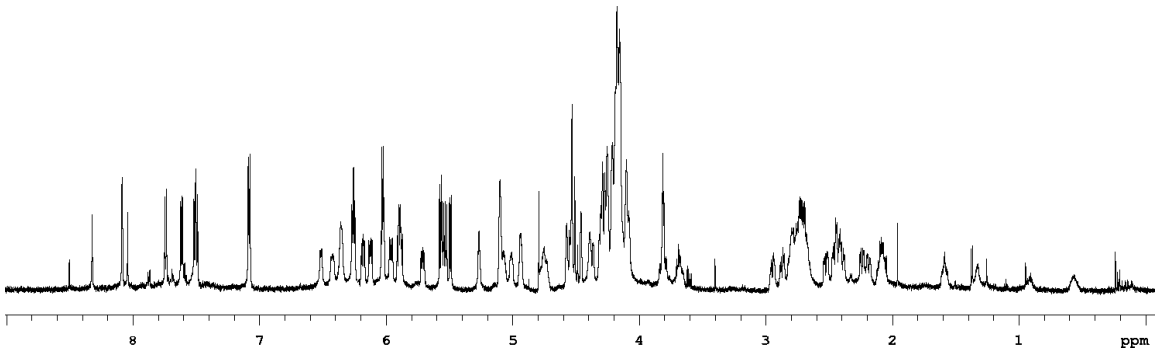
^{1}H-^{1}H NOESY NMR of CG^{8MMPh} in 10 mM Phosphate Buffer pD 7.4 and 500 mM NaCl at 28°C



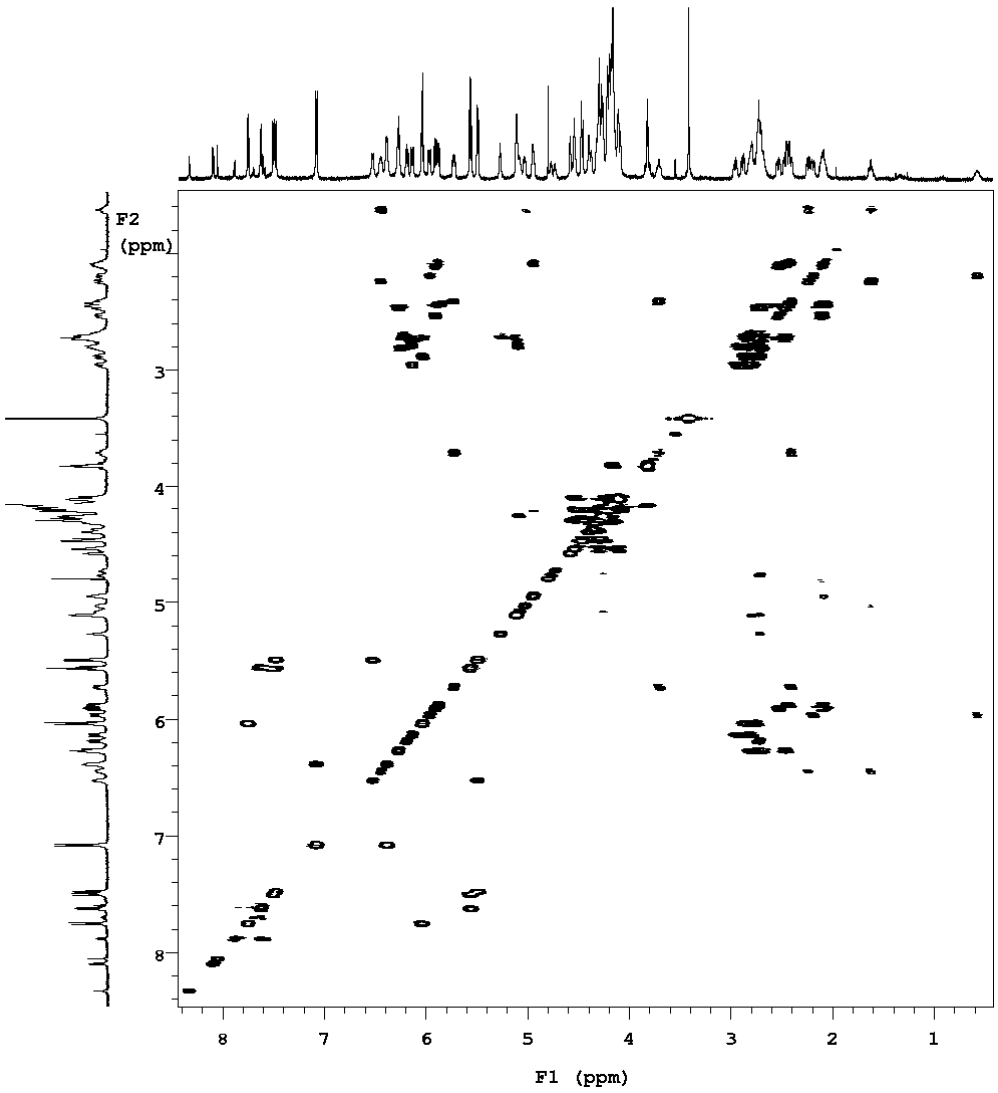
MS^{-} of CG^{8MMPh}



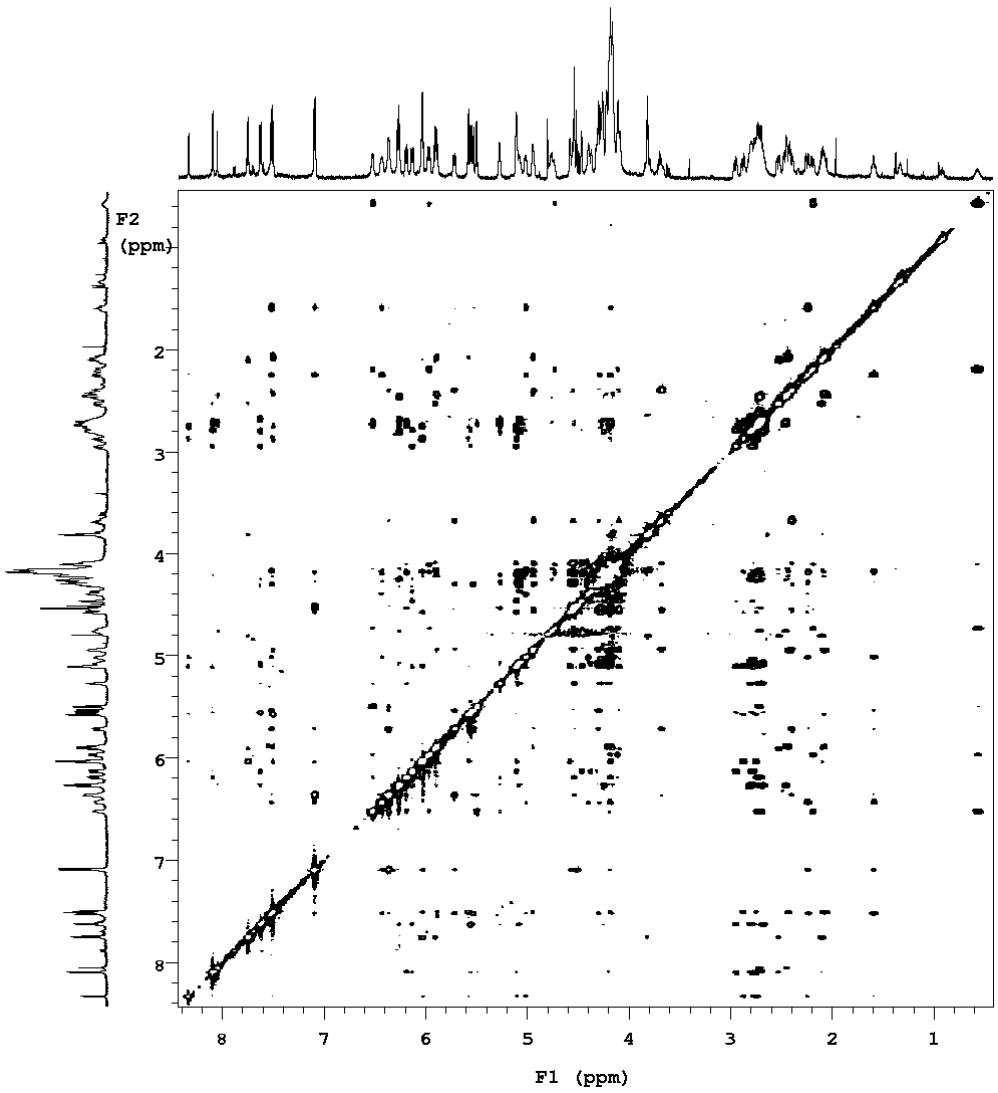
^{1}H NMR of CG^{8HMPh} in 10 mM Phosphate Buffer pD 7.4 at 28°C



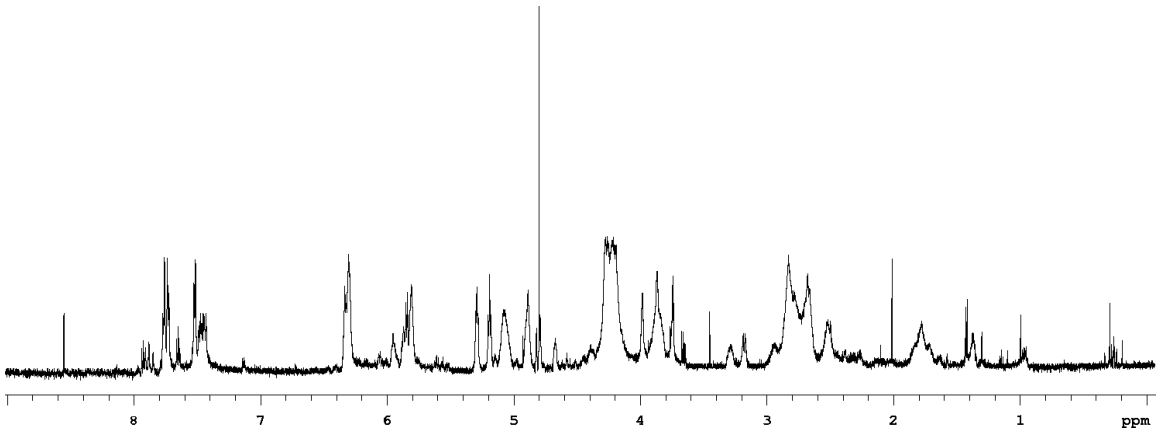
^{1}H-^{1}H COSY NMR of CG^{8HMPh} in 10 mM Phosphate Buffer pD 7.4 at 28°C



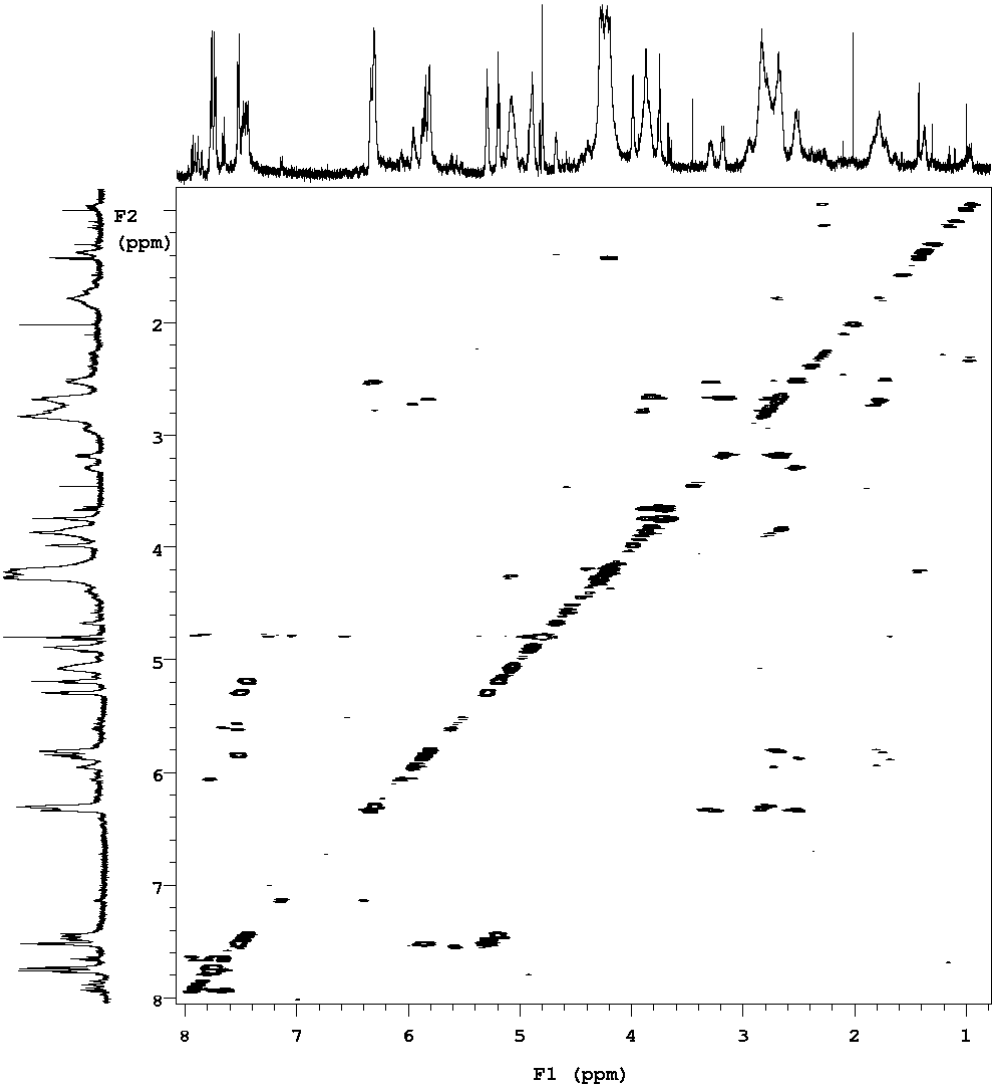
^{1}H-^{1}H NOESY NMR of CG^{8HMPh} in 10 mM Phosphate Buffer pD 7.4 at 28°C



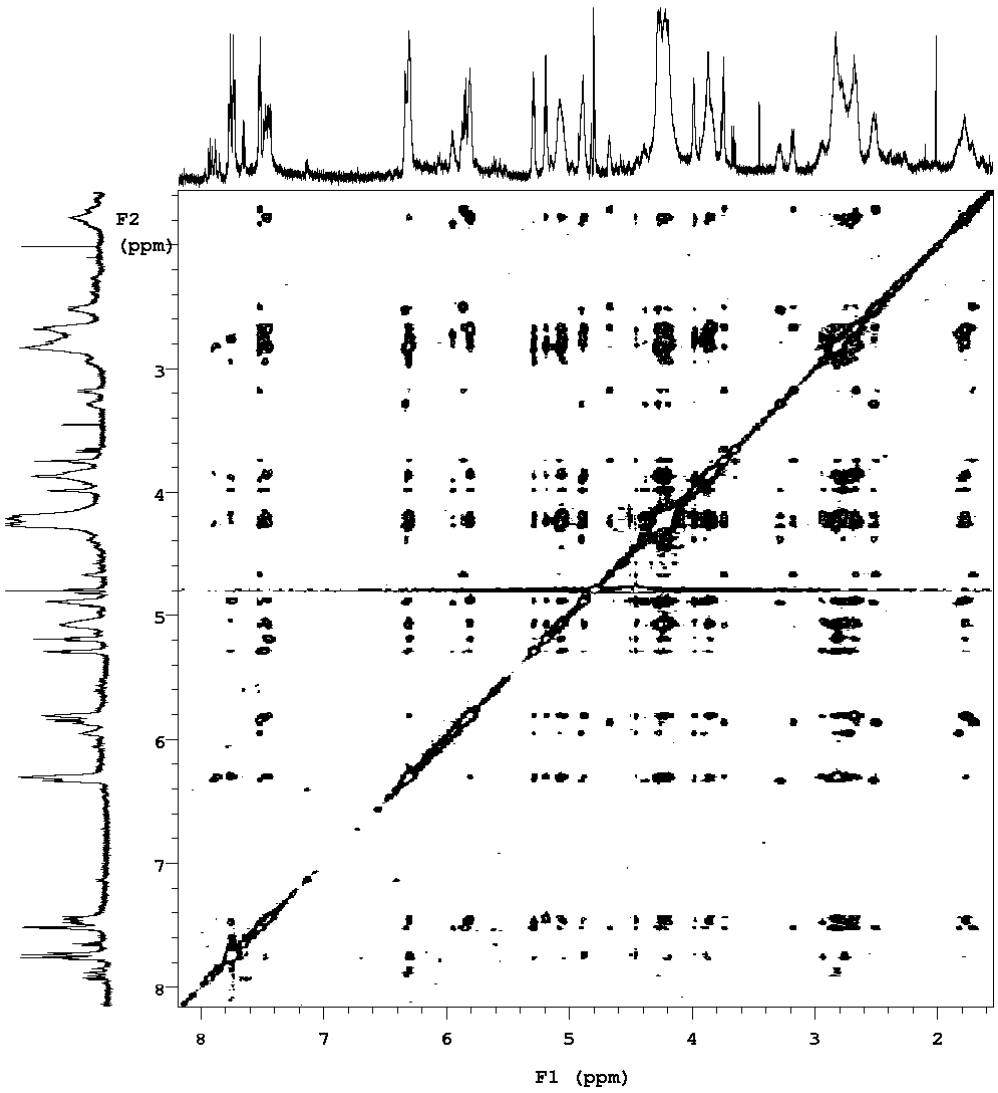
^{1}H NMR of CG^{8HMPh} in 10 mM Phosphate Buffer pD 7.4 and 500 mM NaCl at 28°C



^{1}H-^{1}H COSY NMR of CG^{8HMPh} in 10 mM Phosphate Buffer pD 7.4 and 500 mM NaCl at 28°C



^{1}H-^{1}H NOESY NMR of CG^{8HMPh} in 10 mM Phosphate Buffer pD 7.4 and 500 mM NaCl at 28°C



MS^{-} of CG^{8HMPh}